PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
G01N 27/26, C12Q 1/16
G01N 33/558

A1

(11) International Publication Number: WO 86/ 06835

(43) International Publication Date:
20 November 1986 (20.11.86)

(21) International Application Number: PCT/US86/00984

(22) International Filing Date: 5 May 1986 (05.05.86)

(31) Priority Application Number: 730,618

(32) Priority Date: 6 May 1985 (06.05.85)

(33) Priority Country: US

(71) Applicant: WISCONSIN ALUMNI RESEARCH FOUNDATION [US/US]; 614 North Walnut Street, Madison, WI 53705 (US).

(72) Inventors: DeLUCA, Hector, F. ; 5130 Minocqua Crescent, Madison, WI 53705 (US). BISHOP, Charles, W. ; 424 Milton, Cincinnati, OH 45210 (US). SANTEK, David, A. ; 5166 Old Indian Trail, Madison, WI 53711 (US). KENDRICK, Nancy, C. ; 406 Russell, Madison, WI 53704 (US).

(74) Agent: BREMER, Howard, W.; 614 North Walnut Street, Madison, WI 53705 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: METHOD OF STUDYING GENE EXPRESSION AND OF DIAGNOSING DISEASE

(57) Abstract

A method for studying gene expression and diagnosing disease utilizing a combination of two-dimensional electrophoresis and double label autoradiography for analyzing proteins made as a result of cellular function and which have been separated by the electrophoretic technique. Also disclosed is a fully automatic method for quantitatively analyzing protein spots contained in a developed, double labelled electrophoresis gel.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BG	Bulgaria
BR	Brazil
CF	Central African Republic
CG	Congo
CH	Switzerland
CM	Cameroon
DE	Germany, Federal Republic of
DK	Denmark
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
HU	Hungary
IT	Italy
JP	Japan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
MC	Monaco
MG	Madagascar
ML	Mali
MR	Mauritania
MW	Malawi
NL	Netherlands
NO	Norway
RO	Romania
SD	Sudan
SE	Sweden
SN	Senegal
SU	Soviet Union
TD	Chad
TG	Togo
US	United States of America

Description

METHOD OF STUDYING GENE EXPRESSION AND OF DIAGNOSING DISEASE

Technical Field

The present invention relates to a diagnostic method. More particularly, it relates to a method of discerning and measuring gene expression and diagnosing diseases that are manifested by changes in protein biosynthesis. The measurement and identification of the proteins being synthesized is accomplished using gel electrophoresis and double-label autoradiography coupled with computer analysis of the radiolabel record.

Background Art

There are a number of diseases which produce changes in protein biosynthesis, which changes generally occur before symptoms of the disease become observable. As a result, the detection of changes in protein biosynthesis by cultured cells from a patient offer the opportunity to aid in the diagnosis of such diseases. Such changes also offer the opportunity to study gene expression in cells in general.

In the research laboratory, cellular characteristics are determined by the genes being expressed by those cells as distinctive proteins. Thus the proteins being made are a fingerprint of cellular function and represent the genes being expressed. A method for comprehensively depicting which proteins are being made by certain cells would make available to the scientist a complete picture of which genes are being expressed at any given time. At the present time no such method is available.

Two-dimensional polyacrylamide gel electrophoresis (1) is a powerful technique for separating cellular proteins. Under suitable conditions, each protein contained in a whole cell homogenate or its subcellular fractions can be isolated

within the polyacrylamide gel and visualized by staining or by autoradiography. A finished gel frequently has several hundred to more than 1,000 individual protein spots and provides comprehensive, yet highly detailed, information about cellular proteins. Alterations in protein biosynthesis which are caused by cellular differentiation and transformation by genetic alteration or mutation, or by changes in the cell's environment can be detected by comparing electrophoretic gels containing proteins from experimentally manipulated cells with identically prepared gels containing proteins from control cells. Differences in gene expression between similar types of cells obtained from different sources can also be detected by comparing gels containing proteins from each cell line. However, at the present time detection of these phenomena is limited by the time required to compare precisely the hundreds or thousands of proteins being studied.

Because of the great numbers of proteins which can be involved, computerized systems have been developed to facilitate the comparison of developed electrophoretic gels (2-19). Such systems provide assistance in locating, matching and analyzing spots in gels which represent given proteins. Spot matching can at times, offer problems because of small variations between the electrophoretic separations being compared. Such variations are common and result from differences in batches of ampholines and acrylamide, irregularities in polymerization, and stretching of the isoelectric focusing gel during normal processing. Spot matching can be made more difficult as the result of increased gel to gel variations caused by interfering substances such as lipids, deoxyribonucleic acids or high concentrations of salt.

The aforementioned problems in spot matching can be completely eliminated, however, by the use of double-label autoradiography (19-23). In this process, the proteins in one cell sample is labeled with ^{3}H and mixed with a second or

manipulated cell sample protein labeled with ^{14}C. The combined proteins are then coelectrophoresed to yield a double-labeled two-dimensional gel. Proteins from either sample can be differentially detected by impregnating the gel with a scintillator and exposing it sequentially to two-x-ray films, one sensitive only to ^{14}C and another to both ^{3}H and ^{14}C radiation. Since the images recorded on the paired films are completely superimposable, identically located protein spots can be directly compared.

The process of locating protein spots in images of two-dimensional polyacrylamide gels is difficult to computerize. It is complicated by the great variety of spot shapes, sizes and densities, and by fluctuations or streaks in the film background. It is further confounded by proteins that closely migrate during electrophoresis and touch or overlap in the finished gel. Despite these difficulties, a number of computerized processes have been developed. In some (2,11,17), broad image areas are identified which have densitites greater than an arbitrary threshold, and the positions of individual spots within these areas are located by progressively raising the threshold. In others (4,8), a second derivative of density is computed for each pixel in the image during a single raster pass. Spot locations are then identified by reference to inflections in the resulting secondary image. In still others (6,12,13,18), spots are found by detecting local maxima within the image. Some of these processes automatically match corresponding protein spots that are present in different images, while others rely on operator interaction to various extents. Each computerized process is the product of a new or different approach and is designed to function well for certain applications.

The combination of two-dimensional electrophoresis (1) with various methods for double-label autoradiography (2-6) has already proven to be a valuable method for comparing complex protein mixtures. It has been successfully applied to

the estimation of human genetic polymorphism (7,8), the identification of human gene products (9), the mapping of polypeptide genes with hybrid cells (10), the genetic analysis of human lymphocyte proteins (11), and the detection of a gene product that may relate to Duchenne muscular dystrophy (12). The method also has been used for the detection of gene products of 1,25-dihydroxy-vitamin D_3 (13, 14). Routine application of this powerful method has been limited by the time required by either manual or semi-automatic examination of the great numbers of proteins present in the double-labeled gels.

To date, no fully automatic process has been described for the analysis of double-label autoradiography after electrophoresis. It obviously would be desirable to have a method which is fully automatic and capable of quantitatively analyzing each protein spot contained in a double-labeled gel.

It is an object of the present invention to provide a method for diagnosing diseases which are manifested by changes in cellular protein biosynthesis.

It is a further object to provide a rapid, fully automatic, method for quantitatively analyzing protein spots contained in developed, double-labeled electrophoresis gels.

It is a still further object of this invention to provide a method for studying gene expression.

Disclosure of Invention

The methods of the present invention comprise a comprehensive analysis of cellular protein biosynthesis as evinced by electrophoretic gels obtained from two-dimensional gel electrophoresis, combined with double-label autoradiography.

To study gene expression in cells or tissue: control cells are incubated with ^{14}C leucine; experimental (hormone stimulated or modified) cells are incubated with ^{3}H leucine; the cells are then mixed, fractionated and/or solubilized and subjected to two-dimensional electrophoresis.

For the case of diagnosis of disease: cells obtained from a diagnostic candidate are incubated with ^{3}H leucine; cells from control patients are incubated with ^{14}C leucine; the cells are mixed and subjected to two-dimensional gel electrophoresis.

In both of the above cases the gels are impregnated with a flour, such as Enhance (NEN-Dupont, Boston), and exposed sequentially to a sensitive x-ray film that records ^{3}H and ^{14}C radiation, and to a film that records only ^{14}C. (The gels in each case have included strips with standard ^{14}C and ^{3}H protein for calibration purposes.) The films are then analyzed by:

(a) exactly aligning the x-ray film from the control patient with the
x-ray film from the diagnostic candidate;

(b) scanning and digitizing the optical density of the media containing the protein spots and storing the digitized image as a set of pixel numbers in a first storage means;

(c) calculating the optical density gradient of each pixel and proceeding to (d) when gradient exceeds a preselected threshold;

(d) identifying the center of each protein spot in the image by:

(i) using pixels which have a gradient number exceeding the first threshold as a center, scanning the density of surrounding pixels until the pixel of greatest density is found and repeating the scanning of surrounding pixels using that pixel as the center until the center pixel of greatest density is found;

(ii) storing the center pixel locations in a second storage means;

(iii) calculating the density gradient of pixels in a plurality of directions radially outward from each center pixel; and

(v) storing as a set of boundary locations in a third storage means the position of each pixel which has a density gradient which drops below a second threshold which indicates that the spot ends within the area of said pixel.

In the preferred embodiment the first to fourth storage means are contained within a single computer.

The method is effective and is computationally simple. A pair of images can be processed in approximately 5-min (2.5 min CPU time) thereby greatly reducing the time spent in analyzing experimental data. The preferred method requires 500 kilobytes of memory but can be modified so that it requires only 50 kilobytes with a 10-15-fold increase in processing time. It can also be transported to other systems which access digital image files one line at a time.

The spot location method can locate protein spots whether they are completely resolved within the image or lie within unresolved clusters. Spot boundaries are delineated without shape restrictions and unresolved spots are "split" along an intervening density minimum. Overlapping spots, however, are not resolved mathematically, as in some prior art methods (4,6,8,12,14,27). In cases where the boundaries of spots are inappropriately positioned, the method allows for corrections to be made with operator interaction. Allowance is also made for editing dust spots and other spurious areas of film darkening from the final spot location file stored on a disk. Finally, the method can produce summaries of image analyses in various forms, including hard copy facsimiles of the images which identify the located spots and displays on the video monitor which can be photographed.

The genes being expressed are demonstrated by the ratio of $^{3}H:^{14}C$ in a spot rising above the ratio found in all other spots. The computer can display the protein whose ratio uses above normal values or diminishes below normal thus illustrating whether genes are turned on or off. For diagnostic purposes, the absence of a spot, or the change in

migration of a protein or the addition of a protein would be indicative of a possible physiological defect.

These and other objects and advantages of the invention will be apparent from the drawings and the description of the preferred embodiment.

Brief Description of Drawings

Fig. 1 is a fluorographic image selected for demonstrating the spot locating process. A 10% polyacrylamide gel containing cytosolic proteins labeled with both ^{3}H and ^{14}C (150 mg protein; 1.72×10^{6} ^{3}H dpm; 1.72×10^{5} ^{14}C dpm) was impregnated with EN^{3}HANCE and exposed to a pre-flashed X-Omat AR film for 48 h at -70°C. The molecular weights of the resolved proteins range from 220,000 (top) to 17,000 (bottom), and the isoelectric points range from 8 (left) to 4 (right). The spot identified by the arrow is redisplayed in Fig. 2A.

Fig. 2 shows the relationship between density values and calculated gradients. A well-resolved protein spot from the fluorographic image (Fig. 1) is displayed at full resolution in A. Cross-sectional views of this spot in the direction of the pH gradient are shown for density (B) and density gradient (C) at reduced resolution. Calculation of gradient values for each pixel within the spot's boundaries yields a second degree surface which is shown as a contour plot in D.

Fig. 3 shows the location of a detected spot. Density values within the spot displayed in Fig. 2A are shown in reduced resolution. Bold face numbers designate pixels having calculated gradient values at or above a threshold of 50, and light face numbers designate pixels with sub-threshold of gradient values. An iterative search for the spot's maximum begins with any pixel having a calculated gradient which equals or exceeds the threshold. A search initiated by 48 at the upper left begins in block AI where 145 is identified as the pixel having the greatest density value within the 7 x 7 pixel block (block B) where 209 is identified as the pixel with the greatest density value. Finally, block C, centered

on 209 is searched for greater density values. As no greater values are found, 209 is identified as the spot's maximum, and the image coordinates of this pixel are designated as the spot's location.

Fig. 4 shows spot detection at various gradient thresholds. A complex area of the fluorographic image (Fig. 1) is displayed in A and gradient values in both A and B where values of 50 are identified by solid lines, values of 100 by broken lines, and values of 300 by solid lines labeled 300. Protein spots detected with preselected thresholds of 300, 100, 50, and 25 are marked (±) in C, D, E, and F, respectively.

Fig. 5 shows the accurate location of protein spots amid prominent vertical and horizontal streaks. Most protein spots can be successfully discriminated from streaks by varying the gradient threshold. A threshold value of 100 was selected for this image.

Fig. 6 shows the accurate discrimination of overlapping protein spots. Although the two large protein spots located at the center of this figure were incompletely resolved during electrophoresis, the locations of both spots can be correctly identified. A threshold value of 50 was selected for this image.

Fig. 7 shows the delineation of spot boundaries. Boundaries of protein spots located in Fig. 4D were delineated using a threshold (T*) of 12.

Fig. 8 shows the responses of X-Omat AR and No-Screen films to ^{3}H and ^{14}C. Single-labeled calibration strips were exposed to preflashed X-Omat AR film for 48 hours at -70°C and to No-Screen film for 10 days at 22°C. The resulting film images were digitized by a rotating drum densitometer, and displayed on a McIDAS III video terminal. The film density produced by each calibration strip band was determined by averaging the digital data for a 10 x 10 pixel block within the band's image. Density values produced on X-Omat AR film

by ^{3}H (1/W2/) and ^{14}C (_) are shown in A; those produced on No-Screen film are shown in B (n=8). Smooth curves were calculated from the plotted data by a cubic spline fitting routine. (A pixel equals a 0.01 mm^2 areas of film.)

Fig. 9 shows the linearization of X-Omat AR film responses to ^{3}H and ^{14}C. The spline-fitted curves shown in Fig. 1A were linearized between density values of 50 (---) and 205 (...) by re-expressing the ordinate as shown. Linear regression equations were calculated for the transformed curves with selected data points (1/W2/, ^{3}H:^{14}C) having non-transformed density values f 50, 75, 100, 125, 150, 175 and 200. Both equations, given in the text, had regression coefficients (r) greater than 0.99.

Fig. 10 shows the response of X-Omat AR film to mixtures of ^{3}H and ^{14}C. Calibration strips containing mixtures of ^{3}H- and ^{14}C-labeled proteins were exposed to the same films described in Fig. 8. The densities produced on the X-Omat AR films are plotted against the known ^{3}H content in A. Densities joined by lines were produced by calibration strip bands containing the same amount of ^{14}C (a, 83.3 ^{14}C dpm/pixel; b, 164 ^{14}C dpm/pixel; c 347 ^{14}C dpm/pixel; n=8). The density values ( ) are reploted in B against the total isotopic content of the strips, expressed as ^{3}H alone. The resulting curve closely coincides with the ^{3}H response curve (——) replotted from Fig. 8A.

Fig. 11 shows matching X-Omat AR and No-Screen images produced by exposure t the same double-labeled gel. Cytosolic proteins labeled with both ^{3}H and ^{14}C were electrophoretically resolved within a 10 ; polyacrylamide gel. The gel, containing 150 1/m2/g of protein, 1.72 x 10^{6} ^{3}H dpm, and 2.00 x 10^{5} ^{14}C dpm, was impregnated ith En^{3}HANCE and exposed to an X-omat AR film (A) for 48 h at -70°C, and to a No-Screen film (B) for 10 d at 22°C. Molecular weights of the resolved proteins range from 220,000 (top) to 4 (right). Images produced by a ^{14}C

calibration strip are located to the left of the gel images, and a similar image produced by a ^{3}H calibration strip is located to the right of the gel image on the X-Omat AR film only. The No-Screen film records disintegrations of ^{14}C along by direct autoradiography. The computerized process then calculates the absolute ^{3}H and ^{14}C content of each protein spot from the paired films by knowledge of their responses to the isotopes.

Fig. 12 is a perspective view of a computer system which employes the present invention.

Fig. 13 is a graphic representation of one of the steps performed as part of the present invention.

Fig. 14A and 14B are a flow chart of the processing carried out by the computer system of Fig. 12.

Fig. 15 is a schematic representation of the data structures which are produced by the computer system of Fig. 12 while carrying out the process of Fig. 14.

Description of the Preferred Embodiment

Referring first to Fig. 12, a computer system employed to practice the present invention can be seen which includes an IMB 4381 (Group II) central processing unit 1 which is coupled to a 640 megabyte disc drive 2 (IBM Model 3350) and a 9-track, high speed tape drive unit 3. The processor 1 is accessible through a number of peripheral devices, including a dot matrix printer 4 (Okidata Model 93) and a Televideo CRT terminal 5. An image display monitor 6 (Aydin Controls) is also connected to the processor 1 and it includes local memory for storing a series of 480 by 640 pixel image frames. The stored images may be displayed on its screen 7 and the digital data within any stored image frame can be identified and accessed by using a variable-sized cursor which may be moved on the screen by a joystick 8. As will be described in more detail below, the operator may control the system by entering information through a terminal keyboard 9 in response to prompting messages produced on the terminal display 10.

Matched films produced from the two-dimensional gels are digitized on an Optronics P-1000 densitometer 11 connected to the tape drive 3 by a microprocessor-based controller (not shown). The density of each 0.01 mm^2 film area (pixel) is converted to an 8 bit number ranging from 0 to 255 (0.0 to 3.0 optical density units) which is recorded on magnetic tape. Digitized images (1400 x 1800 pixels) may then be entered into the processor 1 and stored at full resolution in the disc drive 2. These digitized images may be displayed on the monitor screen 7 and they may be processed according to the present invention to automatically locate protein spots.

The computer system of Fig. 12 executes a stored program to process a selected digitized image stored in the disc drive 2. This process is indicated by the flow chart of Figs. 14A and 14B. Referring particularly to Figs. 14 and 15, the selected image is read from the disc drive 2 and is processed by a set of instructions indicated by process block 20 to produce another image that is written out to disc drive 2. This second image of density numbers is more suitable for display on the monitor 6. Each element of the second image is calculated by taking the average value of successive 3-element blocks of the first image. The resulting 467 by 600 element image is a set of 8-bit numbers which each represent the average optical density of a block of pixels in the original digitized image.

As indicated by decision block 21, the operator can begin the protein spot location program. As will become apparent from the description below, the gradient threshold is a variable in the process which enables the operator to control the sensitivity of the protein spot location. By entering smaller values at the keyboard 9 as indicated by process block 21, smaller protein spots will be located. Of course, such increased sensitivity also increases the likelihood that artifacts in the image data will be improperly identified as valid protein spots. Once the gradient threshold value (G_T)

has been set, a loop is entered in which an optical density gradient value is calculated for each pixel, or element, of the averaged image.

As indicated by process block 22, the gradient for each element in the averaged image is calculated by using the values of the eight adjacent optical density numbers. If the D array element is indicated as D(X,Y), then the gradient value G(X,Y) at this pixel is calculated as follows:

$$G(X,Y) = [D(X-1,Y+1)+D(X,Y+1)+D(X+1,Y+1)] - [D(X-1,Y-1)+D(X,Y-1)+D(X+1,Y-1)] + [D(X-1,Y-1)+D(X-1,Y)+D(X-1,Y+1)] - [D(X+1,Y-1)+D(X+1,Y)+D(X+1,Y+1)]$$

For the element D(2m2) for example, the gradient value is:

$$G(2,2) = [D(1,3)+D(2,3)+D(3,3)]-[D(1,1)+D(2,1)+D(3,1)]+D(1,1)+D(1,2)+D(1,3)]-[D(3,1)+D(3,2)+D(3,3)]$$

where D(1,1) D(1,2) D(1,3) are the density values
D(2,1) D(2,2) D(2,3)
D(3,1) D(3,2) D(3,3)

As indicated at decision block 23, the calculated gradient value is then compared with the gradient threshold G_T, and if it is greater, an iterative search is done to find the maximum as indicated at process block 25. This search is initiated by examining the optical density of a 7 by 7 block of elements in the averaged image which are centered on the current element. When the pixel with the highest density is located, it becomes the center of a new 7 by 7 block of elements which are examined. The search ends when the center of the 7 by 7 block is the highest density element. The location of this spot center C_i in the averaged image is stored in a spot center table 27, as indicated by process block 28. The spot center table 27 is scanned after each center is located to make sure the same center is not entered more than once. Because the spot center may encompass an area of several pixels, each new center within two pixels of an existing center is dropped.

After the entire image has been processed, as indicated at decision block 29, the loop is exited. As shown in Fig. 15, the spot center table 27 contains a list of spot centers C_1----C_i along with their locations (X and Y) in the 467 by 600 averaged image.

Referring still to Figs. 14B and 15, a loop is then entered at 30 in which the boundaries for each located spot are found. As indicated by process block 31 and as shown in Fig. 13, this process starts by examining the change in image density at successive locations radially outward from the spot center C_i. The density numbers in the averaged image are employed for this purpose and a gradient number is calculated for successive elements along the eight paths indicated by the vectors 32. The gradient number is merely the change in optical density between the image element D_i and that of the next element D_{i+1} which is radially outward from the spot center C_i.

$$G_i = D_i - D_{i+1}$$

The edge f the spot is defined as the Point P at which the gradient G_i drops below a second preset threshold value "T". The value of T is selected so that the edge is positioned at the first density value D_i which is less than two standard deviations above the local background optical density for well resolved spots, or at the lowest optical density point between two incompletely resolved spots. By searching in eight directions from the spot center C_i, eight points P1-P8 are thus located.

As indicated by process block 33, an interpolation process is then employed to fit a closed curve to the eight points P1-P8. This curve, indicated by the dashed line 34 in Fig. 13 is calculated by a least squares fit to a general equation of the second degree in which no assumptions are made about spot shape. Many such curve fitting routines are known to the art, but in the preferred embodiment a FORTRAN program entitled "CURV" is employed for this purpose. The resulting

second order equation is then employed to locate a series of thirty-six points B1-B36 which are located at successive ten degree intervals around the spot center C_i. The array location (X and Y) of each boundary point B1-B36 is stored in a boundary point table 35.

As each boundary is found, it is drawn on a graphics image which overlays the protein spot image on monitor screen 7. After the entire image has been processed, the operator can edit boundaries which are in error and redraw an outline which is more reasonable using joystick 8. The boundary point table is saved as a file on the disc drive 2.

Since the second autoradiographic image is directly aligned with the first image, the spot finding and curve fitting steps can be skipped. Using tables 27 and 35, the data are read from the second image and then can be compared with the first image.

The following are detailed descriptions of the method of the present invention.

Example 1

Materials and Methods

Software and hardware. Programs for locating, integrating and matching protein spots were written in IBM FORTRAN 77. These programs were run on the Man-computer Interactive Data Access System III (McIDAS III)[2] located within the Space Science and Engineering Center at the University of Wisconsin-Madison. The McIDAS III (25) consists of an IBM 4381 (Group II), twelve 3350 disk units (each capable of storing 640 megabytes), and a printer. The central processing unit is accessible through video terminals which consist of an image display monitor, an alpha-numeric CRT, a keyboard and two cursor control joysticks. Each video terminal has a series of image and graphic frames that are stored locally for rapid access. Image frames (480 x 640 six-bit pixels) can be sequentially displayed and overlayed with graphic frames (480 x 640 three-bit pixels). Digital data within any image frame can be

accessed by the operator with a variable-sized cursor controlled by the joysticks.

Autoradiographs and fluorographs. Cytosolic proteins, labeled with ^{3}H and ^{14}C, were extracted from cultured embryonic chick duodena and resolved by the O'Farrell method of two-dimensional polyacrylamide gel electrophoresis (1). The resulting double-labeled gels were impregnated with scintillator, dried under vacuum and exposed sequentially to two types of x-ray film. One type of film (Kodak X-Omat AR-5) recorded the disintegrations of both isotopes within the gels, and the other type of film (Kodak No-Screen 2T) recorded the disintegrations of ^{14}C alone. The preparation of these matching x-ray film images is described hereinafter.

Digitization of film images. Matching X-Omat AR and No-Screen film images were precisely aligned (± 1.8 mm) during digitization. Good alignment of the images was easily achieved during film exposure by fastening each double-labeled gel to a masonite board having the same dimensions as the x-ray films. When both films were aligned with this board during exposure, the images produced by the gel were positioned in nearly the same place. Small differences in image registration were corrected prior to digitization by overlaying one film on the other and realigning the images with the aid of a light box. When aligned, a single new edge was cut for both films with a straight-edged paper cutter. This new edge was then used to mount identically the matched films on the rotating drum of a densitometer.

Matched films produced from the two-dimensional gels were digitized on a Optronics P-1000 densitometer (Chelmsford, MA) connected to a tape drive via an on-line microprocessor. The density of each 0.01 mm^2 film area (pixel) was converted to an 8 bit number ranging from 0 to 255 (0,0 to 3.0 optical density units) and recorded on magnetic tape. Digitized images (1400 x 1800 pixels) were entered into the McIDAS III for processing and stored at full resolution. Reduced images were created by

replacing each 3 x 3 block of pixels with a single pixel having a grey scale value equal to the mean of the 9 replaced values. These averaged images were used for displaying and interactively examining the digital data on a video terminal and for automatically locating the protein spots.

To begin the spot location process, reduced fluorographic and autoradiographic images of the same two-dimensional gel were entered into a video terminal. The alignment of matching images was visually checked by overlaying the images on the monitor. Images with registration differences in excess of 2 pixels in the horizontal or vertical directions were rescanned before analysis. The area of the gel to be analyzed was then specified by positioning a variable sized cursor box over the appropriate area of the fluorographic image.

To demonstrate the spot finding process, a fluorographic image (Fig. 1) was selected which is representative of the images we have generated. This image contains more than 300 protein spots of many shapes, sizes and densities, and is characterized by the presence of streaks, background fluctuations, and clusters of incompletely resolved spots. The sections which follow focus on selected areas of this image.

The relationship between density and the calculated gradient is illustrated in Fig. 2. Shown in Fig. 2A is a single protein spot selected from a relatively uncrowded area of the fluorographic image. This spot is not representative of all spots in the image but has a shape that is characteristic of well-resolved spots (1). It has an indistinct border that is approximately circular and a density maximum that is slightly off-center in the direction of lower molecular weight. When examined in the direction of the pH gradient, however, the spot's density is symmetrically distributed about the maximum (Fig. 2B). Fig. 2C shows the calculated gradient values for the same pixels whose densities are displayed in Fig. 2B. Low gradient values identify image

areas where the density remains constant, such as in background areas or at the spot's maximum, and high gradient values identify areas where the density differs greatly with small changes in location. Each gradient value, therefore, is analogous to a first derivative of density as a function of location in two dimensions. If the gradient is calculated for each pixel of this protein spot, the second degree surface displayed in Fig. 2D is formed. This surface rises sharply at the spot's edge but decreases in the area of the spot's maximum. The marked changes in this surface at the extreme edges of the spot provide a sensitive means of detecting the spot's boundaries.

To detect protein spots throughout the entire fluorographic image, each pixel's gradient value is compared to a preselected threshold (see below). This comparison is made immediately following each gradient calculation. Pixels having gradient values below the threshold value are predominantly located in the background areas of the image and are ignored. Conversely, pixels with gradient values which equal or exceed the threshold are located within the boundaries of a spot. These pixels signal the presence of a protein spot and initiate a search for its exact location.

Location of detected spots. Whenever a protein spot is detected, an iterative search routine is initiated which locates the spot's maximum. The search begins in a 7 x 7 block of pixels centered on the pixel whose gradient value equals or exceed the threshold. The densities of the 49 pixels within this block are then scanned until the pixel with greatest density is identified. This pixel becomes the center of a new 7 x 7 block which is similarly searched for the pixel of greatest density. As illustrated in Fig. 3, the process is repeated until the pixel at the spot's maximum is identified. This pixel is designated as the spot' location and its image coordinates are kept in memory.

In processing the entire fluorographic image, it frequently happens that a spot's location is identified more than once. This repetition occurs when more than one pixel within the spot's boundaries has a gradient value at or above the selected threshold. Repeated entries are prevented by comparing each new location with those sorted in memory and ignoring any duplicates. This maximum of a detected spot will occasionally encompass an area of several pixels when a spot covers a large area of the image or contains sufficient radioactivity to saturate the film during exposure. In this case, each new location which lies within a radius of 2 pixels from a recorded location is omitted. The final list of spot locations is written to a disk file.

Threshold selection. The detection of protein spots requires the pre-selection of a gradient threshold. Selection of a high threshold results in the detection of only those few spots in which adjacent pixels have greatly differing density values. In contrast, selection of a low threshold leads to the detection of virtually every spot in the image. As larger gradient values are generally associated with larger spots, gradient threshold selection is roughly comparable to specifying the minimum size that a spot must have in order to be detected. The operator can therefore control the sensitivity of the spot detection process by varying the threshold value.

The effect of varying the gradient threshold is illustrated in Fig. 4. Displayed in part A of this figure is one of the more complex areas of the fluorographic image. Superimposed on this image area is a contour plot (Fig. 4B) of the corresponding gradient surface which identifies pixels with gradient values in excess of 50, 100 or 300. A close examination of this contour plot reveals that only the three largest spots contain pixels with gradients of 300 or more. Setting the threshold at 300 therefore results in the detection and location of these spots alone (Fig. 4C). As

progressively lower thresholds are selected, the spot detection process becomes increasingly sensitive: a total of 16 spots are detected with the threshold set at 100 (Fig. 4D), 25 spots are detected with the threshold at 50 (Fig. 4E) and 31 spots are detected with the threshold at 25 (Fig. 4F). Each drop in the threshold value allows smaller protein spots to be detected. The minimum increase in density over background required for spot detection at a gradient threshold of 50 is described hereinafter in the paragraph entitled "Demonstration of Maximum Sensitivity in Spot Detection."

Although lowering the threshold value allows greater sensitivity of spot detection, the injudicious selection of a low theshold can result in sporadic identification of streaks or background fluctuations as protein spots. Consequently, the threshold should be sufficiently low to achieve the highest sensitivity possible in the absence of a significant number of errors. An optimal threshold value can easily be determined for a given image, and then used to process all other images produced under identical conditions from similar protein mixtures. In practice, one image from each series of images is processed at several different threshold values and the located spots are checked on the video terminal. The threshold value yielding the most suitable analysis is then selected and used in processing the remaining images.

Accuracy of spot locations. By careful selection of the gradient threshold, protein spots can be accurately located even in problematic image areas. Shown in Fig. 5 is an area of the fluorographic image where the spots are partically obscured by prominent vertical and horizontal streaks. These streaks, which occur with increasing severity with larger protein loads (1), are commonly encountered in images produced from two-dimensional gels. As demonstrated in this figure, most protein spots can be accurately located even when they appear within steaks which trail from larger spots. The spot location process also retains good accuracy in situations

where two (or more) protein spots overlap due to incomplete resolution (Fig. 6). However, image areas characterized by extremely poor resolution, such as the basic edge of the image, should be avoided during processing.

Infrequent situations arise in which a spot cannot be located, as illustrated in Fig. 7. Although the small spot (a) lying next to one of the largest spots in the fluorographic image can be detected, its location is consistently overlooked. This omission results from low image resolution relative to the 7 x 7 pixel block searched for the spot's maximum. A 7 x 7 pixel block searched for the spot's maximum also includes pixels from the larger spot which have greater density values. Any search which begins with the smaller spot necessarily ends in the location of the neighboring larger spot. A similar situation leading to spot omission arises when a spot (b) forms the shoulder of another larger spot. If a spot of particular interest is omitted during processing, these spacing problems can be overcome by reprocessing the image after reducing it to one-fourth of its original size rather than to the usual one-ninth.

Delineation of spot boundaries. The search for a spot's edges begins with the pixel recorded as the spot's maximum and proceeds along eight lines radiating symmetrically outward. A new gradient (G*) is calculated for each pixel in these search lines by use of a simplified formula:

$$G^* = D_1 - D_2$$

where D_1 and D_2 represent density values at the adjacent pixels which also lie in the search line (D_1 lies nearer to the spot's maximum). The spot's edge is defined as the first pixel encountered in each line which has a gradient value less than a specified threshold (T*). This threshold is selected so that the edge is positioned at the first density value which is less than two standard deviations above local background for well resolved spots or at the lowest density value between incompletely resolved spots. The full boundary

of the spot is then delineated by fitting a closed curve to the eight located pixels. This curve is calculated by a least squares fit to a general equation of the second degree which makes no assumptions about spot shapes. The image coordinates for pixels located at every 10 degrees of arc along the calculated curve are written to a disk file. Calculated curves are deleted if they are hyperbolas or parabolas or if any radius of the curve is less than 2 pixels.

Using a video terminal, the boundaries delineated by this technique are superimposed over the fluorographic image to verify that the slection of T* is correct (Fig. 8). Variations in T* change the areas enclosed by the calculated boundaries. As T* is raised, the spot's edges are located at pixels having density values that significantly exceed local background. As T* is lowered, the spot's edges are located at pixels having density values equal to local background.

Pairing comparable spots in the matching images. After the protein spots in the fluorographic image are fully located, they are paired with comparable spots in the matching autoradiographic image. This pairing process is accomplished by locating a spot in the autoradiographic image which has a maximum density value within 2 pixels of each located spot in the fluorographic image. The boundaries of the newly located spots are then delineated by the same curves calculated for the fluorographic image after appropriate registration adjustments. This pairing process is accurate provided the matching images are precisely aligned before analysis.

Demonstration of Maximum Sensitivity in Spot Detection. A 3 x 3 pixel block in the upper left quadrant of a protein spot can be simulated by the following model:

b	b	xb
b	xb	xyb
xb	xyb	xzb

where bI is the background film density, xI is a percentage increase over background density, and Y and z are different

percentage increases over x. Within this model, the gradient equation presented in "Delineation of spot boundaries" section can be solved for xI, as shown below:

$$\text{Gradient } (g) = xb + xyb + xzb - b + b + xb$$
$$b + b + xb - xb + xyb + xzb$$
$$= xyb + xzb - 2b + 2b - xyb - xzb$$

Rearrangement of the second term yields:

$$g = xyb + xzb - 2b + xyb + xcb - 2b$$
$$= 2\,xyb + xzb - 2b$$
$$= 2b \times y + z - 2$$

Since the density increases at the edge of a spot, x, y and z must be greater than 1, and the equation becomes:

$$g = 2b(x(y + z) - 2)$$

Solving the x:

$$x = ({}^{g}/_{2b} + 2)/y + z)$$

The value of x represents the balance percentage increase over background density required for detection of a spot. If, for example, a gradient (g) of 50 is selected as a threshold, film background equals 30, and y and z are 20 and 30% greater than x, respectively, then:

$$x = 2\left(\frac{50}{30}\right) + 2 \;/\; (1{,}2 + 1{.}3) = 1{.}13$$

In this example, the spot's edge is detected at a location where density increases to 13% above background. Solving for y and z allows the calculation of the density values within the 3 x 3 pixel block. These density values are:

30	30	34
30	34	41
34	41	44

and the edge of the spot is located at the center pixel, 4 density units above background.

Example 2

Materials and Methods

Chemicals, animals, and media. X-Omat-AR-5 x-ray film, No-Screen 2T x-ray film, [1] GBX developer and GBX fixer were

obtained from Eastman Kodak Co. EN^{3}HANCE was purchased from New England Nuclear. L-[U-^{14}C]Leu (336 mCI/mmol) and L-[4,5-^{3}H]Leu (56 Ci/mmol) were obtained from Amersham Corp. in aqueous solutions containing 2% ethanol. Eighteen-day-old chicken embryos (White Leghorn) were obtained from Sunnyside Hatcheries (Oregon, WI) and maintained overnight in a humidified incubator before use. Waymouth's 752/1 medium, penicillin and streptomycin were purchased from GIBCO (Grand Island, NY), and low leucine medium was prepared according to the published formulation for Waymouth's 752/1 medium (3&) by reducing the leucine concentration. Low leucine medium with radiolabel contained ^{3}H-leucine and $^{13=4}$C-leucine at specific activities of 3.17 and 0.317 mCi/l/m2/mol, respectively. All other reagents were of analytical grade.

Film calibration strips. Cytosolic protein, labeled with ^{3}H and/or ^{14}C, was serially diluted into identical PAGE solutions (1) containing 10% acrylamide. These solutions were layered between a pair of glass plates to form 120 x I50 x 0.75 mm slab gels composed of 9 horizontal bands. As soon as each band was poured, its upper surface was overlayed with water until polymerization was complete. This surface was then rinsed with water and gently dried with a length of Whatman filter paper (W. and R. Balston, Eng.) before being overlayed again with the next band. The lowermost band of each gel contained no radiolabel and served as a control for background determinations. Each added layer contained successively greater amounts of radiolabeled protein. After polymerization of the uppermost band, the gels were removed from the plates and fixed for 1 h in an aqueous solution containing 10% (w/v) trichloroacetic acid, 10% (v/v) glacial acetic acid and 30% (v/v/) methanol. Fixed gels were soaked with shaking for 1 h in 65 ml of EN^{3}HANCE and for a second h in water, and were dried on filter paper for 30 min at 80°C using a BioRad gel dryer. All bands of the dried gels were

approximately 1 cm in width and contained a uniform level of radioactivity per unit area.

Identical calibration strips were produced from each dried gel by making one-half inch wide vertical slices through the horizontal bands with a paper cutter. The amount of ^{3}H and/or ^{14}C contained in each band was determined by cutting our replicate polyacrylamide discs with a 7 mm diameter cork borer, burning each disc in a model 306 TriCarb sample oxidizer (Packard Instrument Co., Downers Grove, IL), and measuring the released $^{3}H_2O$ and $^{14}CO_2$ by liquid scintillation spectrometry. The measured radioactivity, adjusted for recovery from the sample oxidizer and for counting efficiency, was expressed as disintegrations/0.01 mm^2 area of gel/unit time.

Exposure of x-ray films to calibration strips. Dried calibration strips were taped face up, by the edges, on smooth-sided masonite boards (8 x 10 x 0.25 in) and exposed to preflashed X-Omat AR-5 x-ray films (8 x 10 in). Preflashing was accomplished under conditions which raised the background film density by 0.15 OD with an electronic flash (Model 283, Vivitar Corp., Santa Monica, CA) fitted with 3 filters (38). The sensitized film surfaces were placed in direct contact with the calibraton strips, and the opposite surfaces were covered with protective yellow paper and a second masonite board. Pairs of boards containing films were stacked between two steel plates (14 x 12 x 0.33 in) connected on all sides by bolts which were evenly tightened until the boards were clamped together with high pressure. The clamped stack was then wrapper in aluminum foil to exclude light and was buried in dry ice for 48 h. Following exposure of the X-Omat AR films, the same calibration stips were repositioned, in similar fashion, on non-preflashed No-Screen 2T x-ray film (8 x 10 in). Exposure of these films was accomplished at 22°C for 10 d.

Film processing. Both types of x-ray films were developed manually using GBX developer and fixer. At the end of the specified exposures, the films were quickly removed from between the steel plates and placed in individual film holders designed for use with vertical film processing baths. X-Omat AR films were developed for 5 min at 20°C, fixed for 3 min and washed for 10 min. No-Screen films were developed for 5 min at 20°C, fixed for 10 min and washed for 30 min.

Preparation of double-labeled protein samples. Duodena from 19-day-old chick embryos were cultured according to a previously described technique (39, 40). The duodena were incubated for the first 18 h in Waymoth's 752/1 medium and in similar medium containing a reduced concentration of leucine (74.41/m2/M) for the remaining 6 h. The low leucine medium was replaced with fresh medium after the 21st h of culturing, and with identical medium containing both ^{3}H-leucine and ^{14}C-leucine after the 22nd h. All media contained penicillin (50 units/ml), streptomycin (50 g/ml) and ethanol (0.1%). Tissue uptake of radiolabeled leucine was determined as previously described (14) and found to be linear during the 2 h labeling period. After 24 h of culturing, the duodena were rinsed in 4°C buffer (50 mM PO_4, 150 mM NaCl, pH 7.4), placed in vials by pairs, and frozen on dry ice. At a later time, the contents of each vial were transferred to a Potter-Elvehjem homogenizer containing an additional 0.5 ml of buffer and homogenized on ice with a motor-driven pestle. Cytosolic extracts were obtained from the 4 homogenates by ultracentrifugation (100,000 x g; 45 min; 4°C) and subsequent recovery of the supernatants. Each extract was analyzed for protein content by the Lowry method (41) and for radiolabel incorporation by trichloroacetic acid analysis (35). Radiolabel incorporation was found to be $11.47 \pm 0.04 \times 10^{6}$ ^{14}C dpm/mg of cytosolic protein.

Double-label two-dimensional gels. Radiolabeled cytosolic extracts were lyophilized and resuspended in sodium dodecyl

sulfate-free lysis buffer (1) to a final concentration of 6 mg/ml. A 25 microliter aliquot of each extract was resolved by two dimensional electrophoresis by the Kendrick Laboratory (Madison, WI). Isoelectric focusing was completed at 400 V for 15 h and 800 V for 45 min, and the molecular weight separation was accomplished in 10% acrylamide gels (120 x 150 0.75 mm). Gels were processed and positioned on x-ray films as described above under "Film Calibration Strips."

Digitization of film images. Films exposed to calibration strips and two-dimensional gels were converted to digital images and entered into the McIDAS II for analysis as previously described.

Calibration of the x-ray films. The responses of the x-ray films to ^{3}H and ^{14}C are determined by the use of calibration strips. These strips, prepared by incorporating radiolabeled proteins into bands of 10% prolyacrylamide gel, contain known amounts of radioactivity per unit area. When calibration strips having the radioactive content listed in Table I are exposed to preflashed X-Omat AR film for 48 h at -70°C, they produce density values like those shown in Fig. 8A. X-Omat AR film density is raised above background when the level of radioactivity exceeds 100 ^{14}C or 500 ^{3}H disintegrations/pixel. Beyond these approximate thresholds, greater film densities are produced by ^{14}C than by equal amounts of ^{3}H, reflecting the higher energy associated with ^{14}C disintegrations. When exposed to No-Screen film for 10 d at 22°C, the calibration strips produce density values like those shown in Fig. 8B. The No-Screen film is completely insensitive to ^{3}H, even at levels approaching 100,000 disintegrations/pixel, and is less sensitive to ^{14}C than the X-Omat AR film.[2]

Smooth curves fitted to the calibration strip data are used to convert density values into isotopic disintegrations. No-Scrren densities are converted directly into ^{14}C disintegrations using a ^{14}C standard curve (Fig. 8B). X-Omat AR densities, however, cannot be directly decoded into the

correct numbers of ^{2}H and ^{14}C disintegrations since identical densities are produced by many different combinations of the two isotopes. Instead, X-Omat AR densities are converted into equivalent numbers of ^{3}H disintegrations using a ^{3}H standard curve (Fig. 8A). These numbers are then decoded into the correct amounts of ^{3}H and ^{14}C using an equation which relates the two isotopes on X-Omat AR film.

Relationship of ^{3}H to ^{14}C on X-Omat AR film. As shown in Fig. 8A, each density value lying within the calibrated response range of the X-Omat AR film can be decoded into disintegrations of either ^{3}H or ^{14}C. A density value of 100, for example, decodes into 5,090 ^{3}H disintegrations or 775 ^{14}C disintegrations, and a density value of 200 decodes into 15,600 ^{3}H disintegrations or 2,210 $Y^{14}C$ disintegrations. It follows, therefore, that a number of $Y^{3}H$ disintegrations can be equated t a different number of ^{14}C disintegrations for each value of film density. In the examples above, 5,090 disintegrations of ^{3}H are equal to 775 disintegrations of ^{14}C with regard to their ability to darken X-Omat AR film, and 15,600 ^{3}H disintegrations are equivalent to 2,210 ^{14}C disintegrations.

An equation which related ^{3}H disintegrations ($^{3}H_{dis}$) to ^{14}C disintegrations ($^{14}C_{dis}$) is derived if X-Omate AR film density (D_{XAR}) is expressed as mathematical functions of the isotopes

$$D_{XAR} = f\ (^{3}H_{dis}) \qquad (1)$$

$$D_{XAR} = f\ (^{14}C_{dis}) \qquad (2)$$

and the two functions are equated by elimination of the variable D_{XAR}. Both functions can be specifically defined by a non-linear regression procedure applied to the calibration strip data. However, this complicated procedure is avoided by linearizing the data and defining the functions by simple linear regression. As shown in Fig. 9, a plot of $(D_{XAR})^{0.05}$ ^{14} vs. log disintegrations is approximately linear

for both ^{3}H and ^{14}C at density values ranging from 50 to 205. Calculation of the best-fit lines for these transformed data defines functions 1 and 2 as:

$$(D_{XAR})^{0.05} = 0.0942 \ (\log {}^{3}H_{dis}) + 0.9102 \quad (3)$$

$$(D_{XAR}(^{0.05} = 0.1010 \ (\log {}^{14}C_{dis}) + 0.9680 \quad (4)$$

By equating expression 3 and 4, an equation is derived which directly relates disintegrations of ^{3}H to disintegrations of ^{14}C:

$$^{3}H_{dis} = 4.108 \ (^{14}C_{dis})^{1.072} \quad (5)$$

This equation allows disintegrations of ^{14}C to be re-expressed as disintegrations of ^{3}H with excellent accuracy. The deviation from linearity seen in Fig. 9 affects the conversion by less than 6%, as shown in Table II.

Calculating the isotopic content of a double-labeled gel. By use of an equation relating to ^{3}H to ^{14}C, the isotopic content of any region of a double-labeled gel is calculated from the images produced on X-Omat AR and No-Screen films. The ^{14}C content is calculated by converting No-Screen density values to C disintegrations and summing the disintegrations within the selected area. This ^{14}C total is divided by the exposure time (min) of the No-Screen film to yield ^{14}C dpm/gel area. Density values within the same area of the X-Omat AR film are converted to equivalent ^{3}H disintegrations, summed and divided by exposure time to yield equivalent ^{3}H dpm/gel area. The true ^{3}H content is then calculated from this number by subtracting the ^{14}C content re-expressed (equation 5) as equivalent ^{3}H dpm/gel area.

To test the method, additional calibration strips were constructed which contained known mixtures of ^{3}H and ^{14}C per unit area. When exposed to the same X-Omat AR films as the single-labeled strips, these double-labeled strips produced the density values shown in Fig. 10A. As expected, X-Omat AR film responded independently to both isotopes. When the level of ^{14}C was held constant, the film density rose with

increasing levels of ^{3}H. Similarly, when the level of ^{3}H was held constant, the film density rose with increasing levels of ^{14}C. Next, the radioactivity in each calibration strip band was re-expressed ^{3}H alone. This was accomplished by converting the known disintegrations of ^{14}C to equivalent disintegrations of ^{3}H (equation 5) and adding them to the known disintegrations of ^{3}H. When these ^{3}H totals, expressed as logarithms, were plotted against the corresponding density values (Fig. 3B), the resulting curve was nearly identical to the standard curve for ^{3}H alone (Fig. 8A). This finding demonstrates that X-Omat AR film responds to mixtures of ^{3}H and ^{14}C in the same way that it responds to equivalent amounts of ^{3}H. Finally, the double-labeled calibration strips were exposed to the same No-Screen films as the single labeled stips. Using the method described above, the radioactivity contained in each band was calculated from the images produced on the X-Omat AR and No-Screen films. As determined from Table II, the calculations had errors of 4.0 ± 1.0% for ^{14}C, 9.8 ± 8.1% for ^{3}H, and 8.6 ± 10.0% for the ^{3}H ^{14}C ratio.

Automatic analysis of double-labeled sample gels.

Calculations for determining the isotopic content of double-labeled gels were computerized and further tested with 4 radiolabeled protein samples resolved electrophoretically within two-dimensional gels. Each sample was prepared by culturing embryonic chick duodena in medium containing both ^{3}H- and ^{14}C-leucine and extracting the cytosolic proteins. Since the proteins within each sample were labeled with both isotopes simultaneously, they all contained the same ratio of ^{3}H to ^{14}C. This ratio was found to be 8.64 ± 0.09 by direct analysis. The computerized analysis of these sample gels could therefore be considered correct if a ratio of 8.64 was determined for each spot.

Sequential exposure of the gels to X-Omat AR and No-Screen films produced matching pairs of images like those displayed in Fig. 11. Standard curves for ^{3}H and ^{14}C were

calculated from the density values produced on these films by newly prepared calibration strips. The curves were similar to those shown in Fig. 8 and were entered into McIDAS III for decoding the images into isotopic disintegrations. Protein spots were located as previously described, within the entire gel image excepting the unresolved sections along the edges. Spots having density values greater than the calibrated film response ranges were automatically deleted, and those resulting from digitization artifacts (dust, scratches, etc.) were manually edited from memory. As shown in Table IV, the automatic analysis calculated the istopic content of approximately 250 spots in each gel. The ratio of ^{3}H:^{14}C within these spots averaged 8.07 to 9.03, closely agreeing with the experimentally determined ration of 8.64.

The results of these tests demonstrate that the isotopic content of a double-labeled gel is accurately determined by use of the computerized process. This process, however, yields accurate data only when five experimental conditions are met. First, the x-ray film responses to ^{3}H and ^{14}C, and the equation relating ^{3}H to ^{14}C on the X-Omat AR film, must be determined for each set of films. The shapes of the X-Omat AR response curves change with small variations in preflashing conditions and exposure time (38). Response curves for both types of film also vary with the concentrations of scintillator and acrylamide within the gels, the thickness of the gels, and the conditions of film exposure and development (21, 38, 42). Because of these variables, it is difficult to reproduce the same response curves for different sets of films. Second, calibration strips and sample gels must be processed in identical fashion. Fluorography is reduced by increasing concentrations of acrylamide and protein stain (42, 43), by incomplete saturation of the gel with scintillator, and by increasing thickness of the gel (21). Consequently, calibration strips and sample gels must have the same thicknesses, contain the same acrylamide concentrations, and

be fixed, stained, impregnated with scintillator and dried in identical fashion. Better accuracy is achieved when protein staining is avoided. If EN^{3}HANCE is used, strips and gels should be dried for as short a time as possible and immediately exposed to X-Omat AR films to reduce the loss of the volatile fluor. Third, calibration stips and sample gels must contact the x-ray films with high pressure during exposure to maximize the efficiency with which the radioactive disintegrations darken the films. Sufficient pressure to maintain good contact is easily achieved by using steel pressure plates. Fourth, each set of films (X-Omat AR or No-Screen) should undergo exposure and subsequent development as a single batch. This increases the reproducibility between films by controlling for temperature changes that affect exposure and development (44,45). Finally, exposures must be selected so that imageareas of interest have density values that lie within the calibrated response ranges of the films. This can be accomplished by adjusting the levels of radioactivity within the calibration strips to match the levels of radioactivity contained by the sample, and then selecting appropriate exposure times. When a sample contains a complex mixture of proteins, it is unlikely that an exposure time can be selected which will be appropriate for every protein. With short exposures, radio-labeled proteins which are present in lesser amounts may produce no images at all. With longer exposures, these minor constitutents of the protein mixture may produce detectable images while the major consituents saturate the film with isotopic disintegrations. In such cases, a comprehensive analysis of the protein mixture can only be accomplished by: 1) making a series of identical sample gels and exposing them to films for different lengths of time, or 2) making a series of gels which contain different amounts of the same sample and exposing all of them to films for the same length of time.

As shown in Table III, the computerized process calculated the ^{14}C content of the double-labeled calibration strips with greater accuracy than the ^{3}H content. Calculated ^{14}C values differed from the known values by an average of 4.0 · while calculated ^{3}H values differed from the known values by an average of 9.8%. Preliminary research in this laboratory [4] indicates that the higher error associated with ^{3}H calculations related to problems with fluorography. Irregularities in the gel surface, which frequently appear during drying, decrease the contact between the gel and film, and thereby decrease the efficiency with which radio-active disintegrations are recorded. Formation of these surface irregularities can be greatly reduced by the complete elimination of air bubbles between the gel's surface and the rigid polyurethane cover sheet after placing the gel on the dryer. Fluorography is also disrupted by the incomplete impregnation of sample gels with scintillator. Gels containing a uniform mixture of ^{3}H and ^{14}C per unit area may produce slightly mottled images on X-Omat AR film when impregnated with either EN^{3}HANCE or 2,5-diphenyloxazole (44). This mottling is not apparent in images produced on No-Screen film by the same gels, indicating that the cause related to a heterogeneous distribution of scintillator within the gel. Although its exact cause is unknown, this mottling also contributes to the error in the ^{3}H calculations, and thereby increases the variability of the ^{3}H: ^{14}C ratios calculated for protein spots in the double-labeled sample gels.

The computerized process is designed to facilitate the comparison of complex protein mixtures. By rapidly calculating ^{3}H:^{14}C ratios for large numbers of protein spots contained in a double-labeled two-dimensional gel, it can aid an investigator in detecting differences between paired samples. Such differences are indicated by isotope ratios which differ significantly from a mean ratio calculated for all protein spots detected in a gel. As spot ratios within

the test gels (Table IV) had coefficients of variation (standard deviation/mean) of approximately 25%, proteins whose sysnthesis rate is altered by 50% should be detectable at a significance level of 0.05%. This approach greatly facilitates the analysis of double-label studies that attempt to identify proteins biosynthesized in response to experimental manipulation.

Direct exposure DEF-5 x-ray film, manufactured by Kodak may be used as a suitable replacement for No-Screen 2T x-ray film in double-label autoradiography. The No-Screen film is more sensitive to ^{14}C than X-Omat AR film when both films are darkened by direct autoradiography. In addition to the isotopes employed, other combinations that might be used are ^{3}H and ^{35}S or ^{75}Se and ^{35}S.

TABLE I

Single-labeled calibration strips for determining x-ray film responses.

Multibanded calibration strips were formed from acrylamide solutions containing serially diluted ^{3}H- or ^{14}C-labeled proteins as described under "Materials and Methods." The radioactive content of these strips is shown here.

Band No.	Disintegrations per minute (dpm) x 10^{3}/pixel[a]	
	^{14}C strips	^{3}H strips
1	1200	5800
2	600	2930
3	300	1480
4	149	744
5	74.5	375
6	37.2	189
7	18.6	95.3
8	9.3	48.1
9	0.0	0.0

[a]A pixel is defined as a unit of area having dimension of 0.1 x 0.1 mm.

TABLE II

Accuracy of an equation relating ^{3}H to ^{14}C on X-Omat AR film.

An equation relating the abilities of ^{3}H and ^{14}C to darken X-Omat AR film was derived from the linearized film responses. By use of this equation, ^{14}C disintegrations producing densities within the calibrated response range of the film are converted to an equivalent number of ^{3}H disintegrations with little error.

Digitized film density	Disintegrations/pixel/48 h			% Error[a]
	Known ^{14}C	^{3}H equivalent of ^{14}C	Known ^{3}H	
40	143	841	892	5.72
50	280	1730	1750	1.09
75	536	3470	3390	2.39
100	775	5150	5090	1.01
125	1010	6870	6930	0.85
150	1290	8920	9120	2.26
175	1670	11700	11900	0.02
200	2210	15800	15500	1.91

[a]Calculated as (^{3}H equivalent of ^{14}C)/Known ^{3}H - 1 x 100%

TABLE III

The known and calculated isotopic contents of double-labeled calibration strips.

Double-labeled calibration strips composed of 9 polyacrylamide bands were exposed to X-Omat AR film for 48 h at -70°C, and to No-Screen film for 10 d at 22°C. The ^{3}H and ^{14}C contained in each band was calculated from the film images by reference to standard curves. Calculated values for ^{14}C, ^{3}H and the $^{3}H:^{14}C$ ratio are displayed here alongside the known values.

Band No.	Disintegrations per minute x 10^{-3}/pixel				$^{3}H:^{14}C$ ratio	
	Calculated ^{14}C	Known ^{14}C	Calculated ^{3}H	Known ^{3}H	Calculated	Known
1	83.3	86.8	747	604	8.97	6.96
2	83.3	86.8	1220	1210	14.6	13.9
3	83.3	86.8	2200	2430	26.4	28.0
4	164	169	1400	1210	8.54	7.16
5	164	169	2280	2430	13.9	14.4
6	164	169	4650	4860	28.4	28.8
7	347	329	2500	2430	7.20	7.39
8	347	329	4060	4860	11.7	14.8
9	0	0	0	0	-	-

TABLE IV

Summary of the automatic analysis of double-labeled sample gels.

Gel No.	Spots Analyzed	Ratio of ^{3}H:^{14}C spots (x ± S.D.) Calculated	Known
1	268	8.07 ± 1.40	8.72 ± 0.16
2	278	8.75 ± 2.09	8.71 ± 0.17
3	245	9.13 ± 1.65	8.52 ± 0.06
4	223	8.75 ± 0.44	8.64 ± 0.09
Mean ± S.D.	254 ± 25	8.68 ± 0.44	8.64 ± 0.09

The method of the present invention can be used to detect changes in protein synthesis or gene expression induced by hormones, various agents such as mutagens, recombinant DNA experiments or changes in bathing fluid. It would also be applicable to detection of disease such as genetic defects or viral infection.

References

1. O'Farrell, P. H. (1975) J. Biol. Chem. 250, 4007-4021.
2. Bossinger, J., Miller, M. J., Vo, K. P., Geiduschek, E. P. and Xuong, N.H. (1979) J. Biol. Chem. 254, 7986-7998.
3. Capel, M., Redman, B. and Bourque, D. P. (1979) Anal. Biochem. 97, 210-228.
4. Garrels, J. I. (1979(J. Biol. Chem. 254, 7961-7977.
5. Alexander, A., Cullen, B., Emigholz, K., Norgard, M. V. and Monahan, J. J. (1980) Anal. Biochem. 103, 176-183.
6. Anderson, N. L. Taylor, Jr., Skandora, A. E., Coulter, B.P. and Anderson, N. G. (1981) Clin. Chem. 27.
7. Aycock, B. F., Weil, D.E., Sinicropi, D. V. and McIlwain, D. L. (1981) comput. Biomed. Res. 14 314-326.
8. Lemkin, P. F. and Lipkin, L. E. (1981) Comput. Biomed. Res. 14, 272-297.
9. Lemkin, P. F. and Lipkin, L. E. (1981) Comput. Biomed. Res. 14, 355-380.
10. Lemkin, P. F. and Liplin, L. E. (1981) Comput. Biomed. Res. 14, 407-446.
11. Vo, K. P., Miller, M. J., Geiduschek, E.P., Nielson, C., Olson, A. and Xuong, N. H. (1981) Anal. Biochem. 112, 258-271.
12. Brown, W. T. and Ezer, A. (1982) Clin. Chem. 28, 1041-1044.
13. Fox, S. H. (1982) Clin. Chem. 28, 932-934.
14. Garrison, J. C. and Johnson, M. L. (1982) J. Biol. Chem. 257, 13144-13149.
15. Kendrick, N. C., Eakle, J. and DeLuca, H.F. (1982) Electrophoresis 3, 346-353.
16. Mariash, C. N., Seelig, S. and Oppenheimer, J. H. (1982) Anal. Biochem. 121, 388-394.
17. Miller, M. J., Vo, P. K., Nielson, C., Geiduschek, E. P. and Xuong, N. H. (1982) Clin. Chem. 28, 867-875.
18. Skolnick, M. M., Sternberg, S. R. and Neel, J. V. (1982) Clin. Chem. 28, 969-978.

19. Goldman, R. C., Trus, B. L. and Lieve, L. (1983) Eur. J. Biochem. 131, 473-480.
20. Gruenstein, E. I. and Pollard, A. L. (1976) Anal. Biochem. 76, 452-457.
21. McConkey, E. H. (1979) Anal. Biochem. 96, 39-44.
22. Choo, K. H., Cotton, R. G. H. and Danks, D. M. (1980) Anal. Biochem. 103, 33-38.
23. Finger, J. M. and Choo, K. H. (1981) Biochem. J. 193, 371-374.
24. Bishop, C. W., Kenrick, N. C. Santek, D. A., Thompson, R. G. and DeLuca, H. F. (1984) J. Biol. Chem. (submitted).
25. Suomi, V. E., Fox, R., Limaye, S. S. and Smith, W. L. (1983) J. Clim. Appl. Meteor. 22, 766-778.
26. Rosenfeld, A. and Kak, A. C. (1976) Academic Press: New York, pp. 284-286.
27. Agard, D. A., Steinberg, R. A. and Stroud, R. M. (1981) Anal. Biochem. 111, 257-268.
28. Lecocq, R. E., Hepburn, A. and Lamy, F. (1981) Anal. Biochem. 127, 293-299.
29. McConkey, E. H., Taylor, B. J. and Phan, D. (1980) Proc. Natl. Acad. Sci. USA 76, 6500-6504.
30. Walton, K. E., Styer, D. and Gruenstein, E. I. (1979) J. Biol. Chem. 254, 7951-7960.
31. McConkey, E. H. (1980) Somat. Cell Genet. 6, 139-147.
32. Taggart, R. T. and Franke, U. (1982) Cytogenet. Cell Genet, 32, 99-110.
33. Hamaguchi, H., Ohta, A., Mukai, R., Yabe, T. and Yamada, M. (1981) Hum. Genet. 59, 215-220.
34. Rosenmann, E., Kreis, C., Thompson, R. G., Dobbs, M., Hamerton, J. L. and Wrogemann, K. (1982) Nature 298, 563-565.
35. Bishop, C. W., Kendrick, N. C. and DeLuca, H. R. (1983) J. Biol. Chem. 258, 1305-1310.

36. Santek, D. A., Bishop, C. W., Kendrick, N. C., Thompson, R. G. and DeLuca, H. F. (1984) J. Biol. Chem. (submitted).

37. Waymouth, C. (1959) J. Natl. Cancer Inst. 22, 1003-1017.

38. Laskey, R. A. and Mills, A. D. (1975) Eur. J. Biochem. 56, 335-341.

39. Corradino, R. A. (1973) J. Cell Biol. 58, 64-78.

40. Franceschi, R. T. and DeLuca, H. F. (1981) J. Biol. Chem. 256, 3840-3847.

41. Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951) J. Biol. Chem. 193, 265-275.

42. Harding, C. R. and Scott, I. R. (1983) Anal. Biochem. 129, 371-376.

43. Higgins, C. R. and Dahmus, M. E. (1982) Electrophoresis 3, 214-221.

44. Bonner, W. M. and Laskey, R. A. (1974) Eur. J. Biochem. 46, 83-88.

45. Randerath, K. (1970) Anal. Biochem. 34, 188-205.

Claims

1. A method of locating protein spots in an electrophoresis gel and detecting the boundaries of said spots, said method comprising:
   (a) dividing the area of X-ray film which has been exposed to the gel into a plurality of small area pixels;
   (b) scanning and digitizing the optical density of the film containing a record of the protein spots and storing the digitized image as a set of pixel numbers in a first storage means;
   (c) calculating the optical density gradients of each pixel and storing the resulting set of pixel gradient numbers in a second storage means;
   (d) identifying the center of each protein spot in the image by:
      (i) scanning the set of pixel gradient numbers and locating those pixels in which the gradient number exceeds a first threshold which indicates that the pixel is within a protein spot;
      (ii) using pixels which have a gradient number exceeding the first threshold as a center, scanning the density gradients of surrounding pixels until the pixel of greatest density is found and repeating the scanning of surrounding pixels using that pixel as the center until the center pixel of greatest density is found;
      (iii) storing the center pixel locations in a third storage means;
      (iv) calculating the density gradient of pixels in a plurality of directions radially outward from each center pixel; and

(v) storing as a set of boundary locations in a fourth storage means the position of each pixel which has a density gradient which drops below a second threshold which indicates that the spot ends within the area of said pixel.

2. A method for determining gene expression by comparative electrophoretic analysis which comprises

incubating control cells in radioactively labeled amino acids wherein the radioactive label is characterized by strong radiation emission

incubating experimental cells in radioactively labeled amino acids wherein the radioactive label is characterized by weak radiation emission

mixing the cells

solulizing the mixed cells

subjecting the solubilized cells to two-dimensional gel electrophoresis

impregnating the electrophoretic gel with a fluor

exposing the thus impregnated gel to a first film which is sensitive to both the weak and strong radiation

exposing the thus impregnated gel to a second film which is sensitive only to the strong radiation

accurately aligning the films in superposition

detecting the radiolabeled protein spots in accordance with the procedure of claim 1 and

computing the amount and ratio of the radiolabels (isotopes) in each spot

whereby the particular genes expressed or suppressed in the experimental cells can be identified.

3. The method of claim 2 wherein the strong radiation emitter is ^{14}C and the weak radiation emitter is ^{3}H.

4. The method of claim 3 wherein the strong radiation emitter is ^{35}S and the weak radiation emitter is ^{3}H.

5. The method of claim 3 wherein the strong radiation emitter is ^{75}Se and the weak radiation emitter is ^{35}S.

6. A method for diagnosing a disease in a patient which is characterized by changes in protein biosynthesis which comprises

taking a skin biopsy of the patient

culturing the fibroblasts from said skin

incubating said fibroblasts in ^{3}H leucine

incubating fibroblasts from a normal subject in ^{14}C leucine

mixing the resulting ^{14}C and ^{3}H labeled cells

subjecting the mixed cells to two-dimensional electrophoresis

impregnating the electrophoretic gel with a fluor

exposing the thus impregnated gel to a first film which is sensitive to both the weak and strong radiation

exposing the thus impregnated gel to a second film which

is sensitive only to the strong radiation

accurately aligning the films in superposition

detecting the radiolabeled protein spots in accordance with the procedure of claim 1 and

computing the amount and ratio of the radiolabels (isotopes) in each spot

whereby alterations in protein biosynthesis associated with certain disease states can be identified.

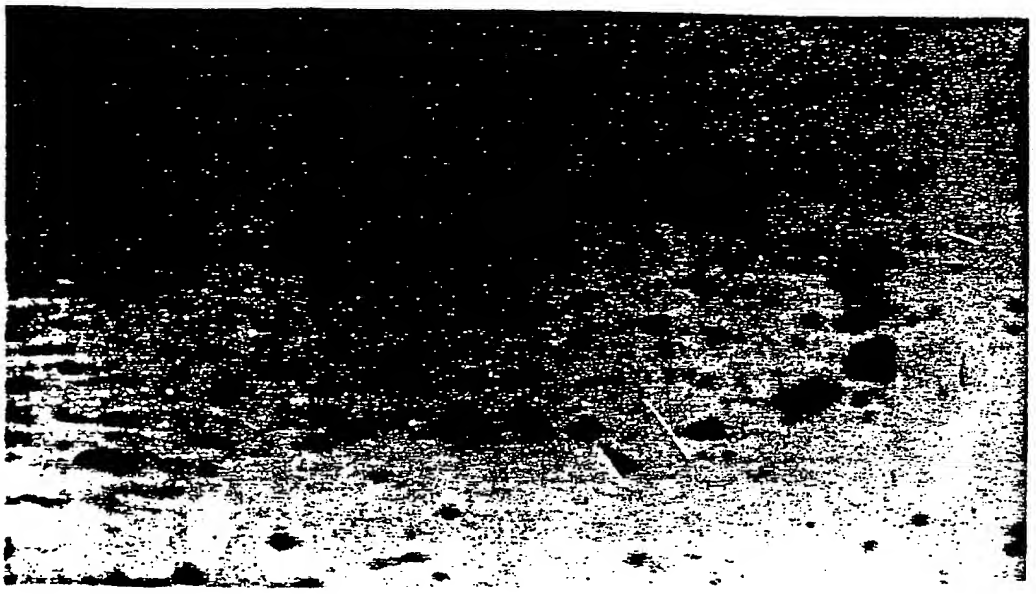

FIG. 1

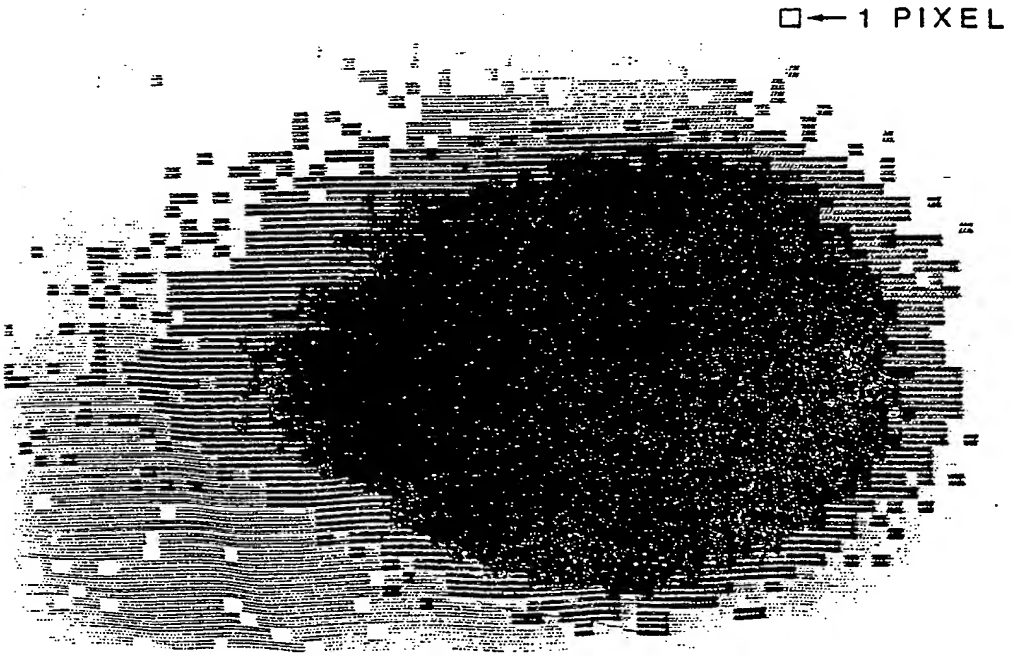

FIG. 2A

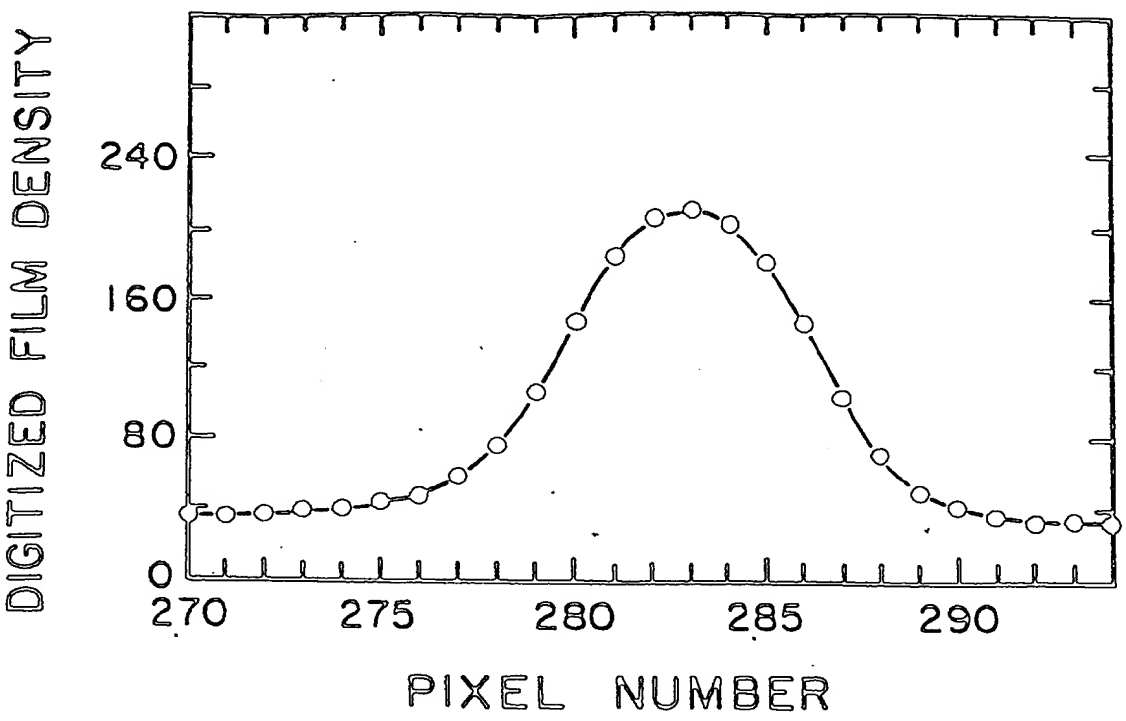

FIG. 2B

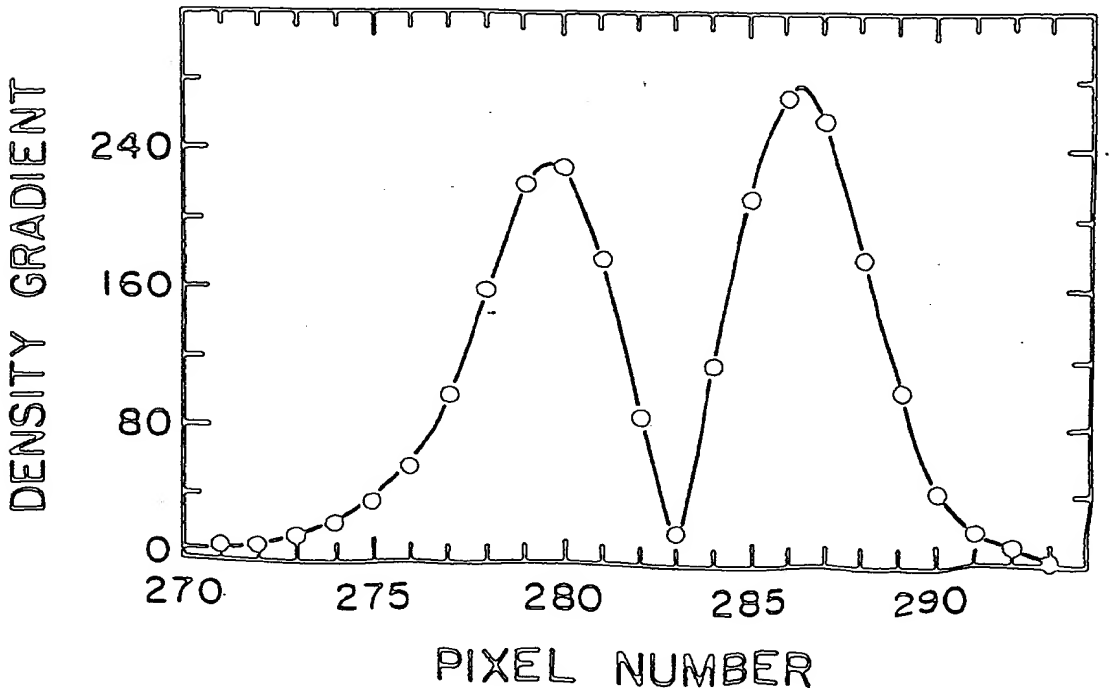

FIG. 2C

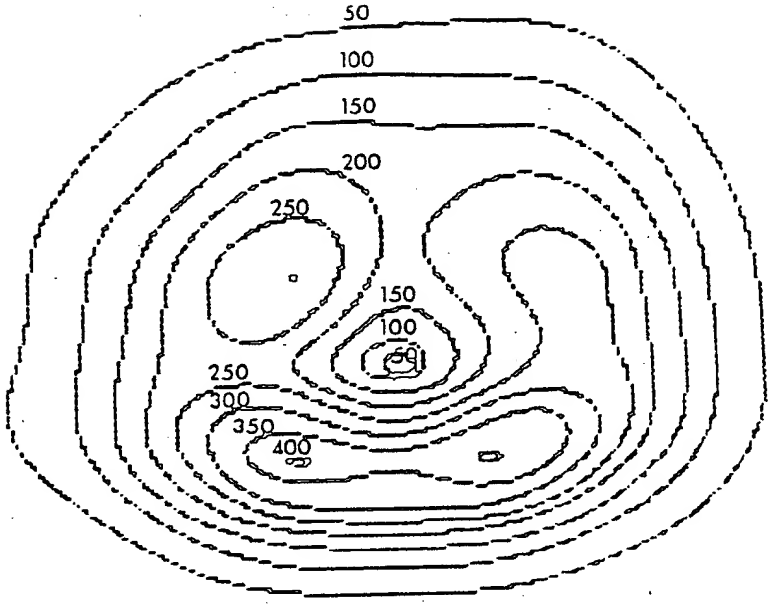

FIG. 2D

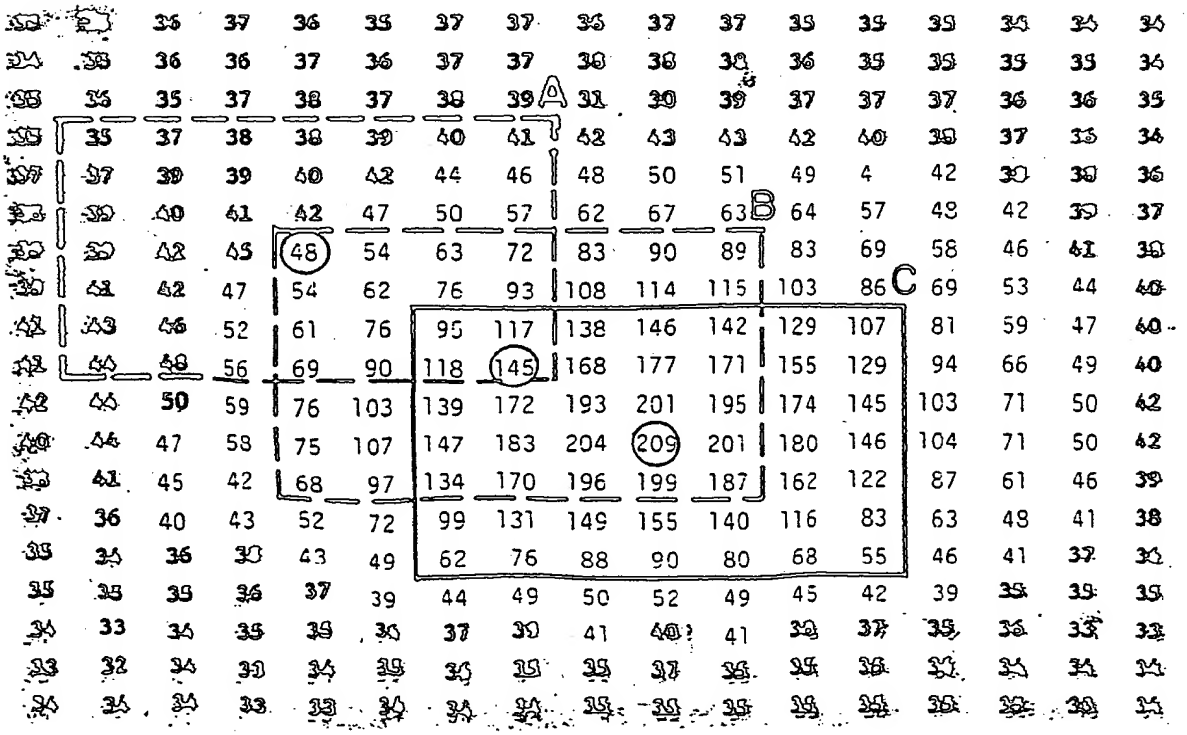

FIG. 3

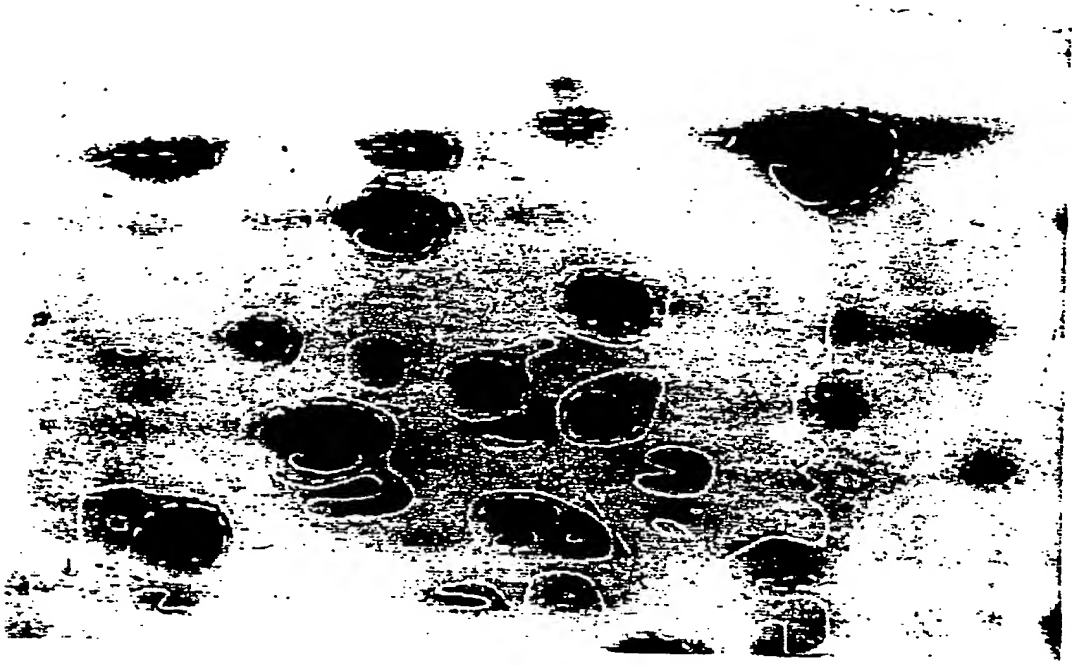

FIG. 4A

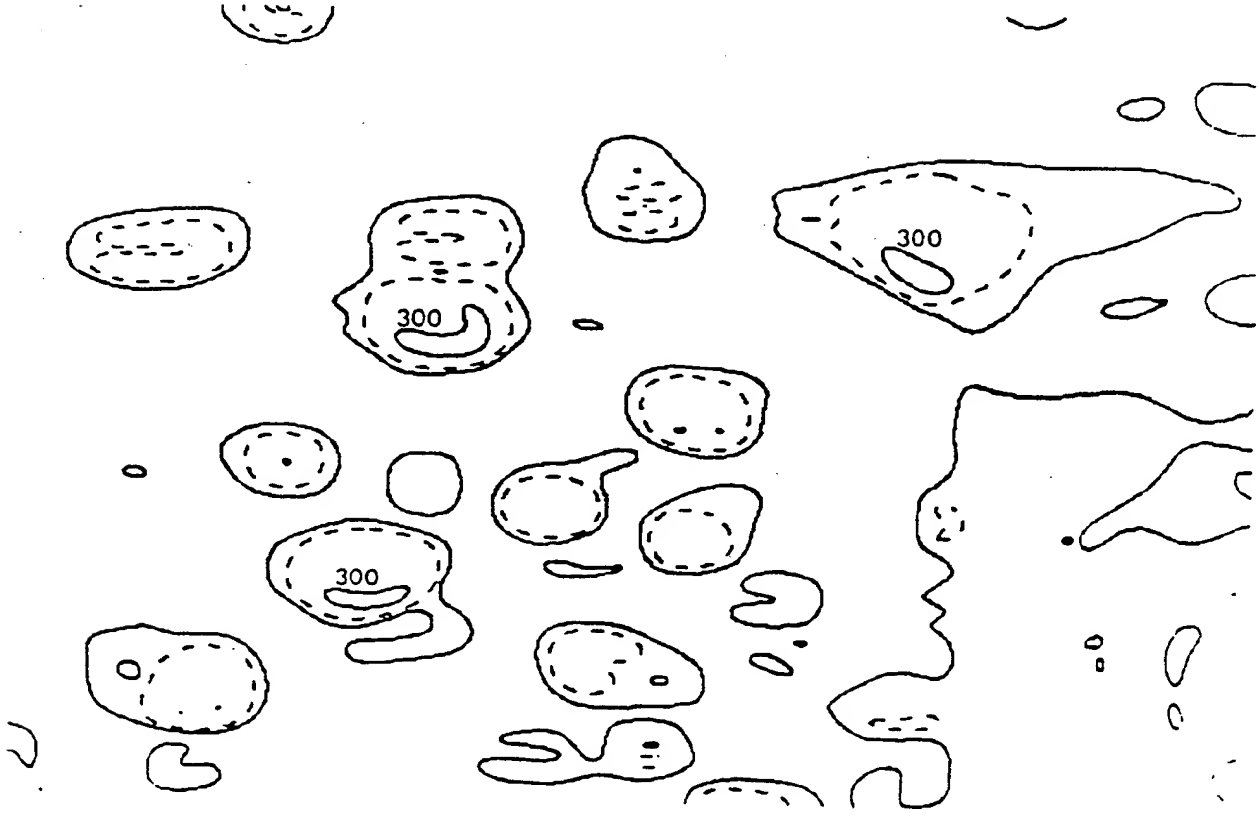

FIG. 4B

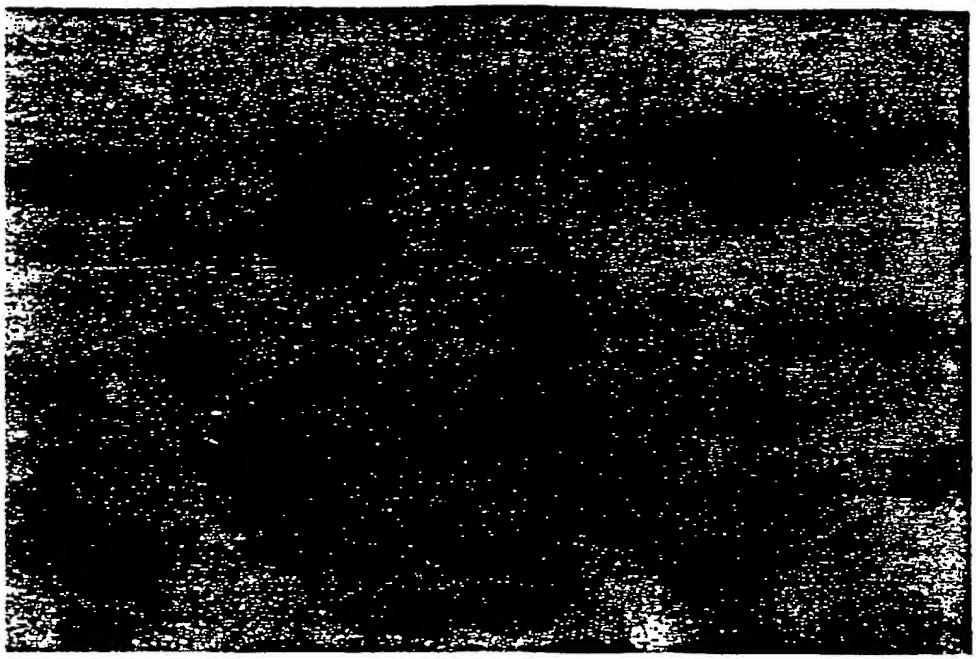

FIG. 4C

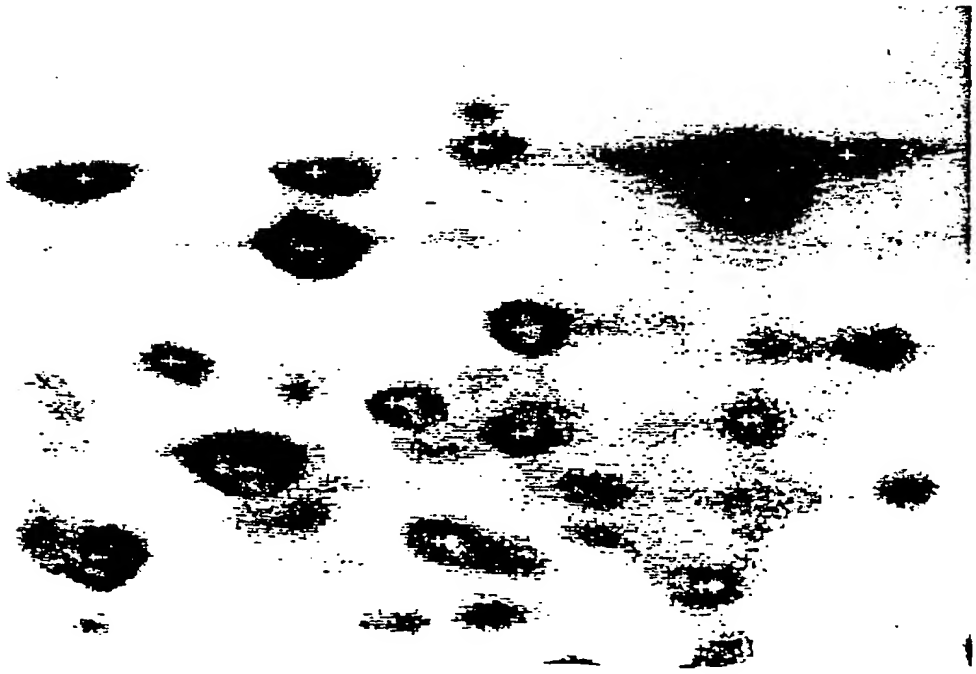

FIG. 4D

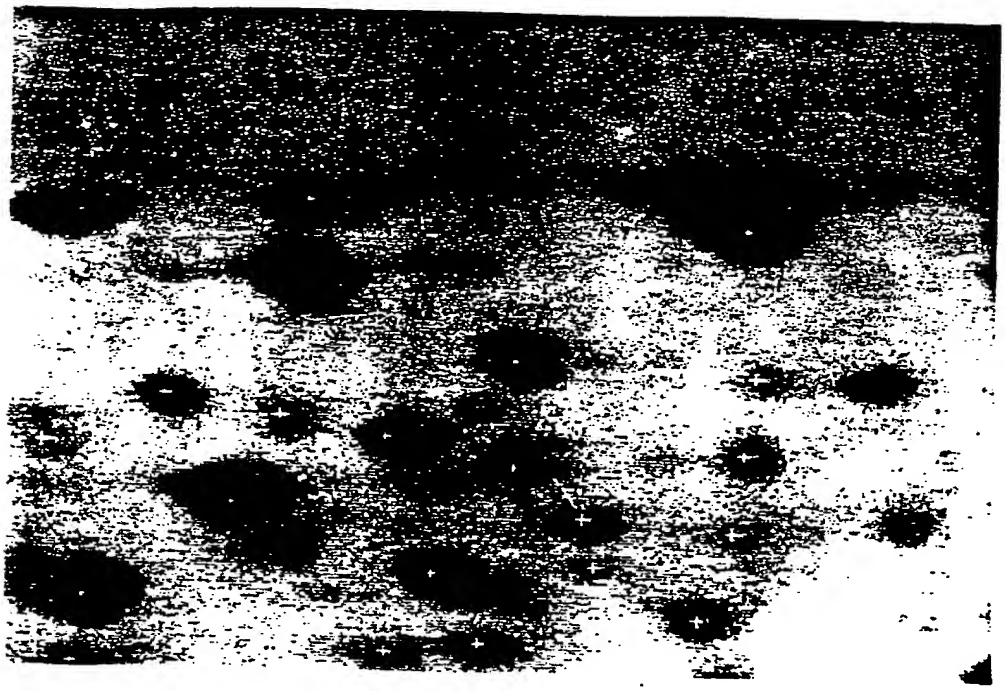

FIG. 4E

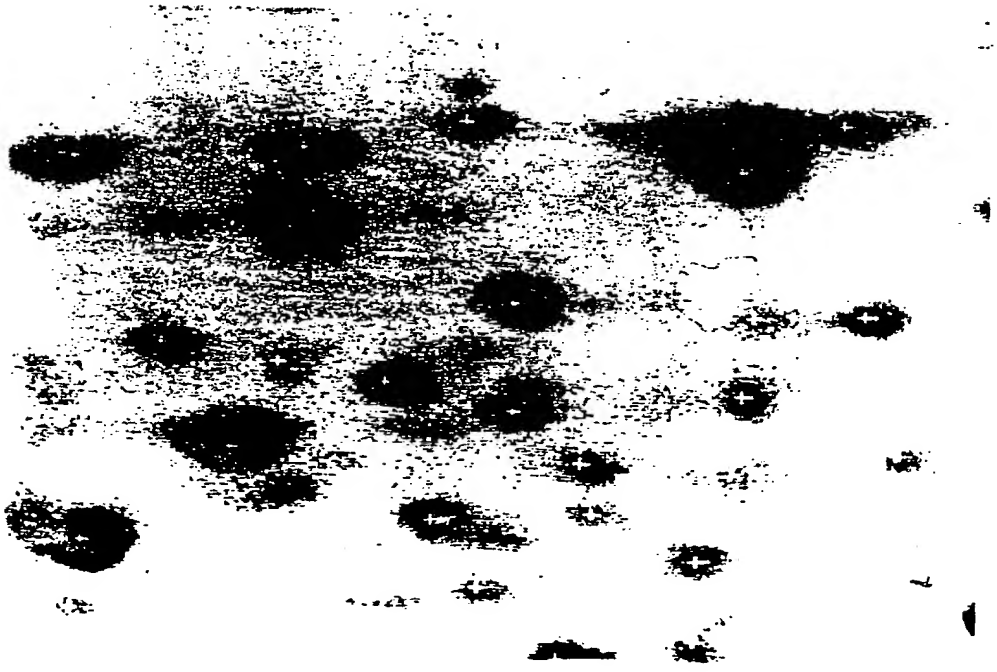

FIG. 4F

FIG.5

FIG.6

FIG.7

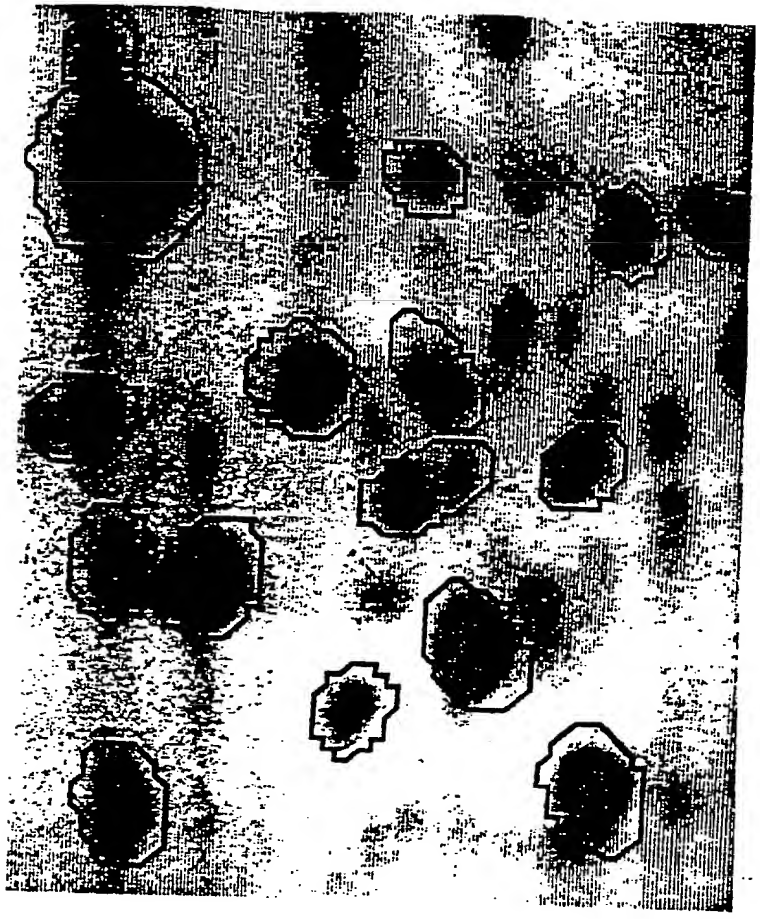

FIG.11A

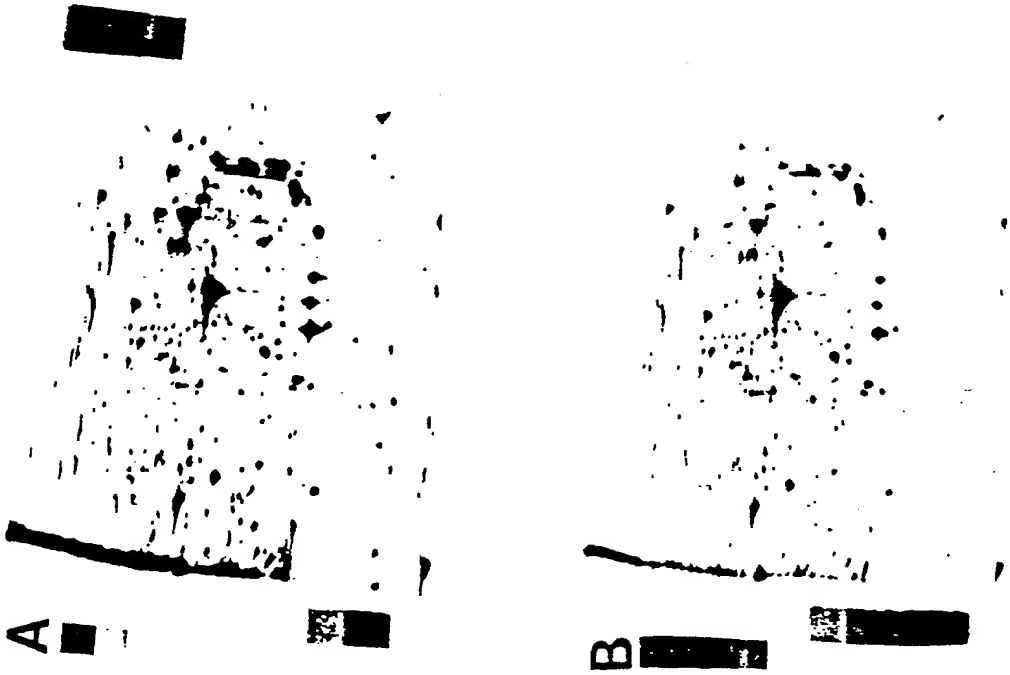

FIG.11B

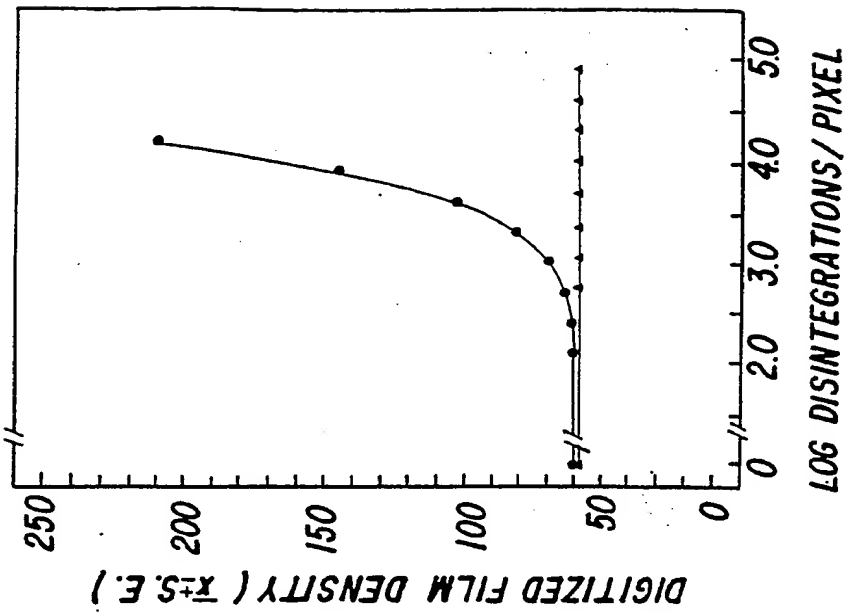

Fig.8A

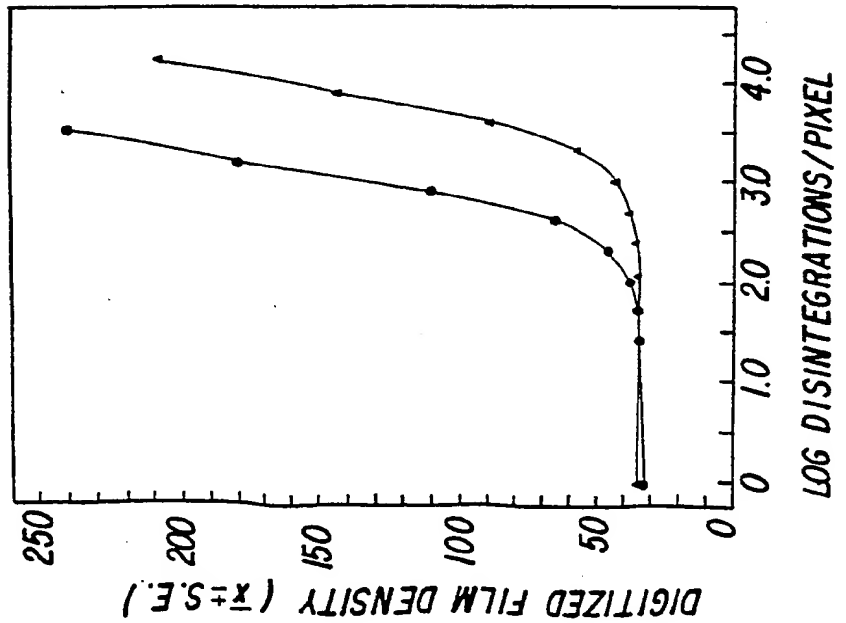

Fig.8B

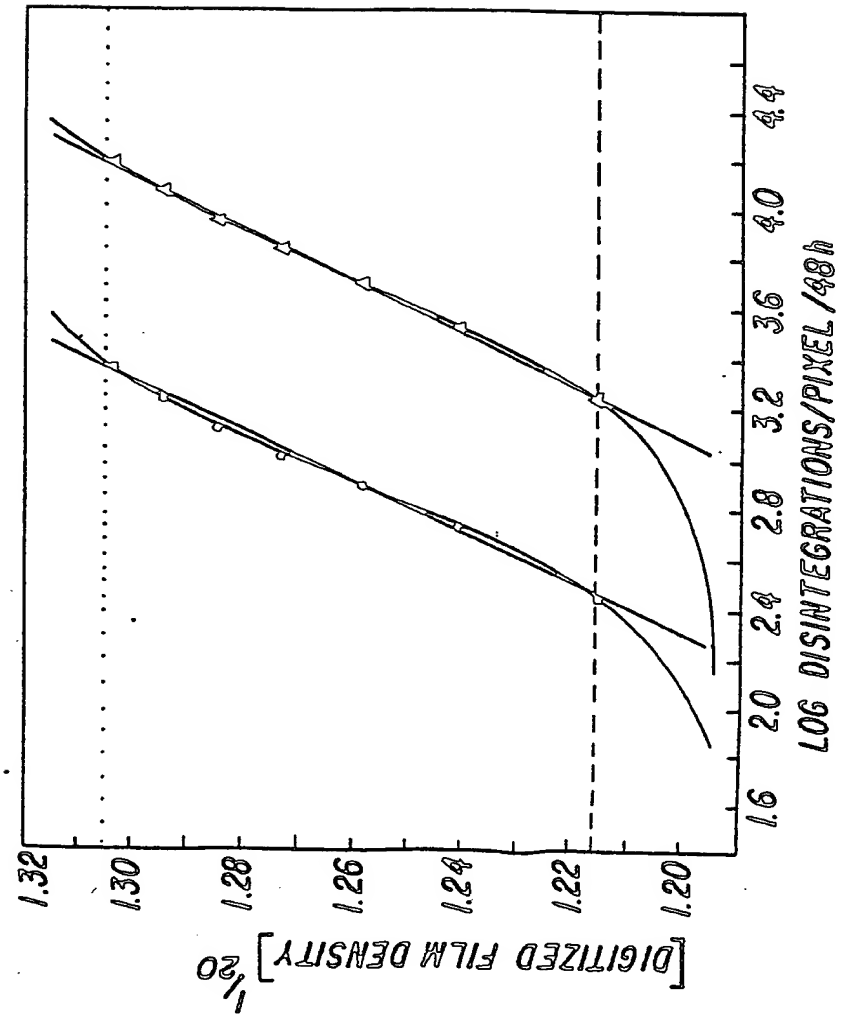

Fig. 9

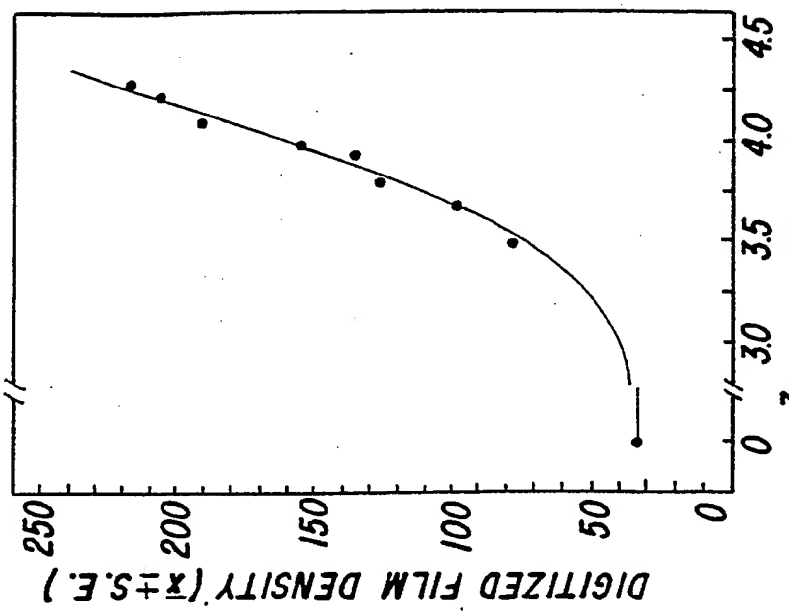

Fig. 10B

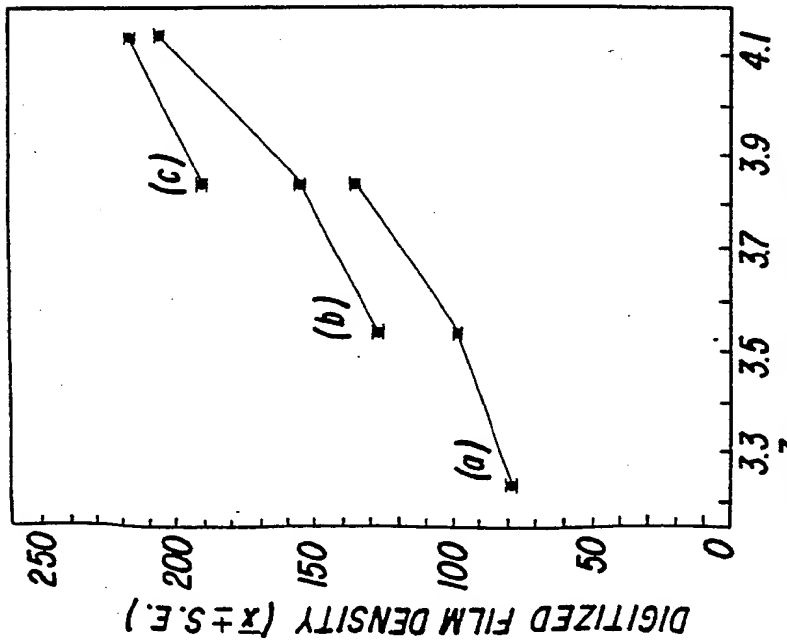

Fig. 10A

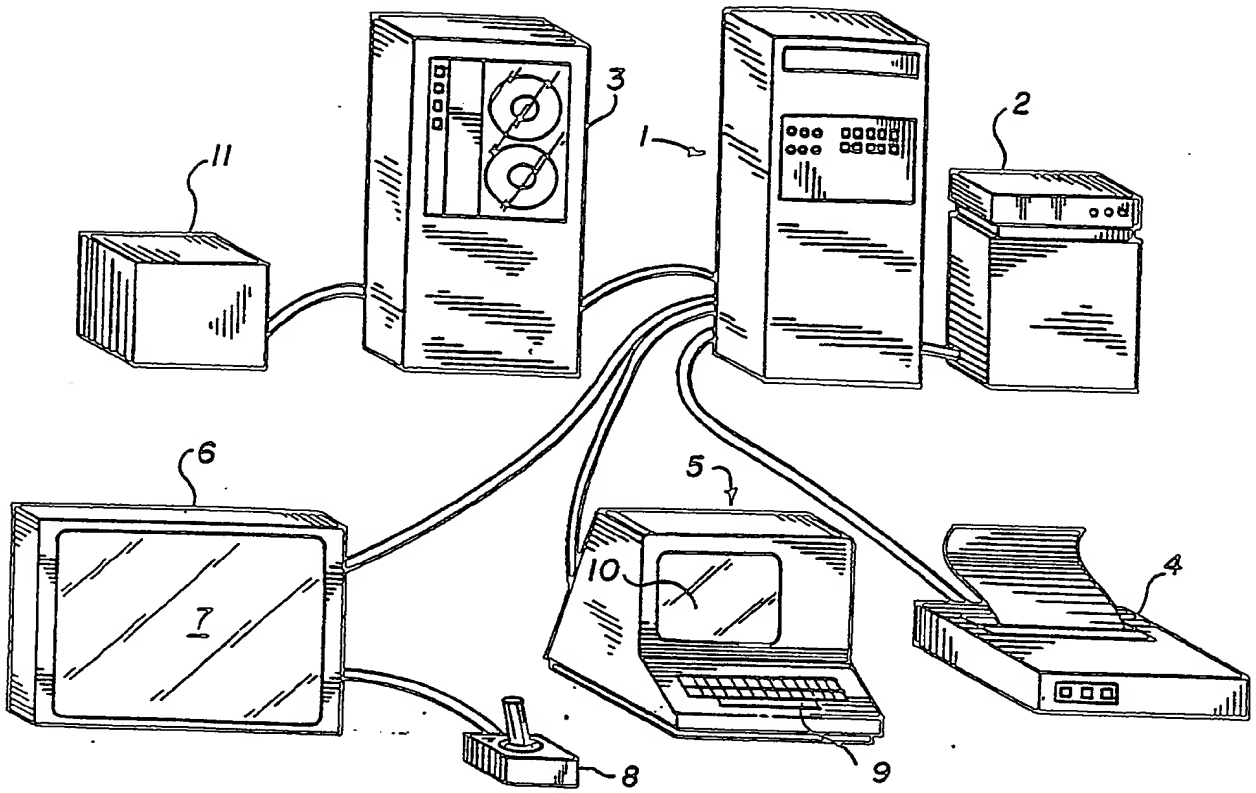

FIG. 12

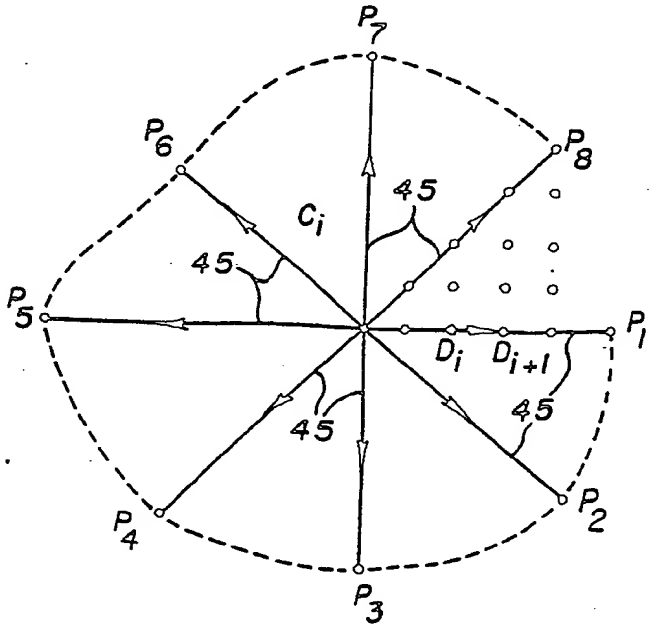

FIG. 13

SUBSTITUTE SHEET

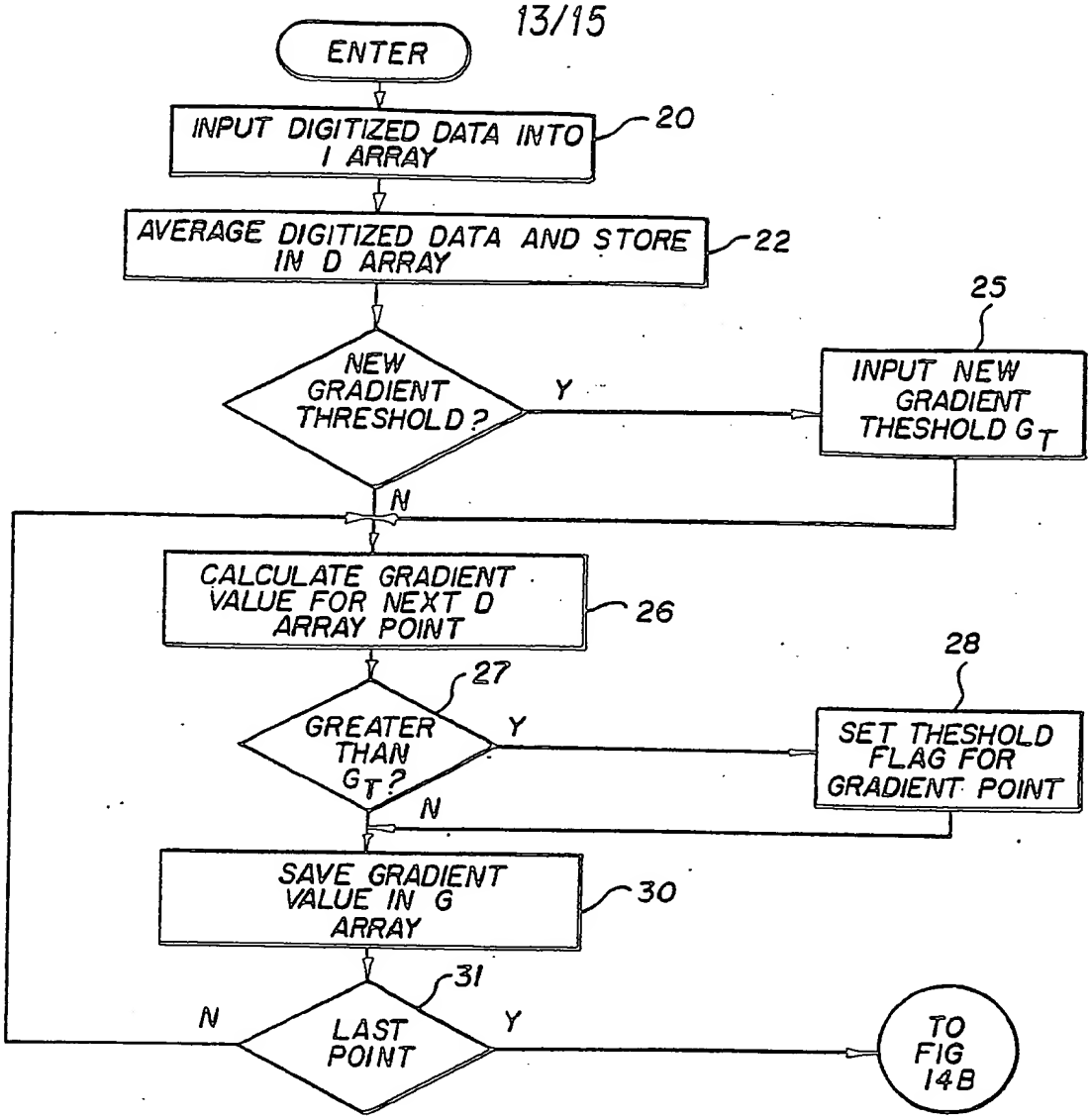
ENTER
INPUT DIGITIZED DATA INTO I ARRAY
20
AVERAGE DIGITIZED DATA AND STORE IN D ARRAY
22
NEW GRADIENT THRESHOLD ?
Y
25
INPUT NEW GRADIENT THESHOLD G_T
N
CALCULATE GRADIENT VALUE FOR NEXT D ARRAY POINT
26
27
GREATER THAN G_T ?
Y
28
SET THESHOLD FLAG FOR GRADIENT POINT
N
SAVE GRADIENT VALUE IN G ARRAY
30
31
N
LAST POINT
Y
TO FIG 14B

FIG. 14A

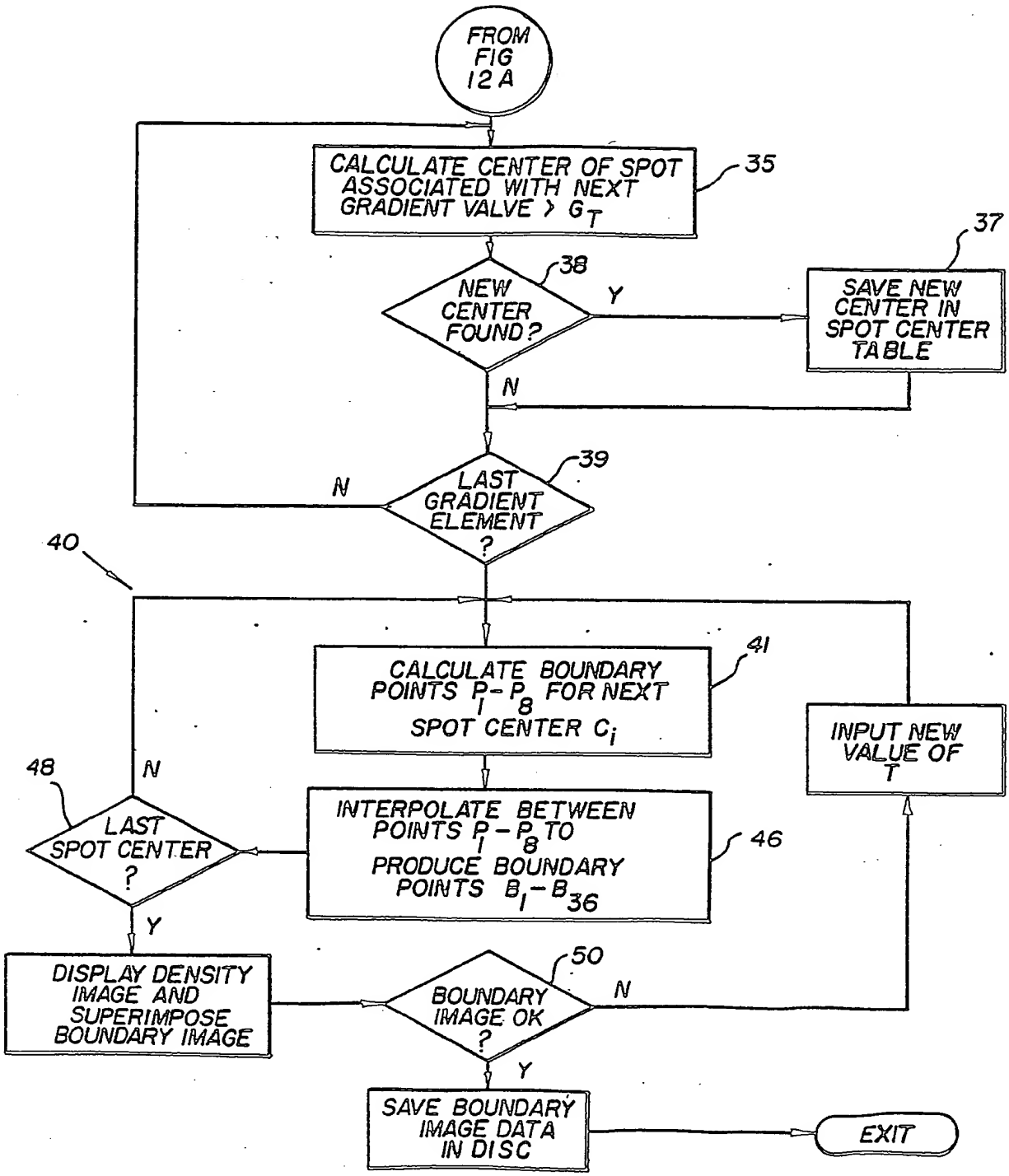
FROM FIG 12A
CALCULATE CENTER OF SPOT ASSOCIATED WITH NEXT GRADIENT VALVE > G_T
35
NEW CENTER FOUND?
38
Y
37
SAVE NEW CENTER IN SPOT CENTER TABLE
N
LAST GRADIENT ELEMENT ?
39
N
40
CALCULATE BOUNDARY POINTS P_1-P_8 FOR NEXT SPOT CENTER C_i
41
INPUT NEW VALUE OF T
48
N
LAST SPOT CENTER ?
INTERPOLATE BETWEEN POINTS P_1-P_8 TO PRODUCE BOUNDARY POINTS B_1-B_{36}
46
Y
DISPLAY DENSITY IMAGE AND SUPERIMPOSE BOUNDARY IMAGE
BOUNDARY IMAGE OK ?
50
N
Y
SAVE BOUNDARY IMAGE DATA IN DISC
EXIT

FIG. 14B

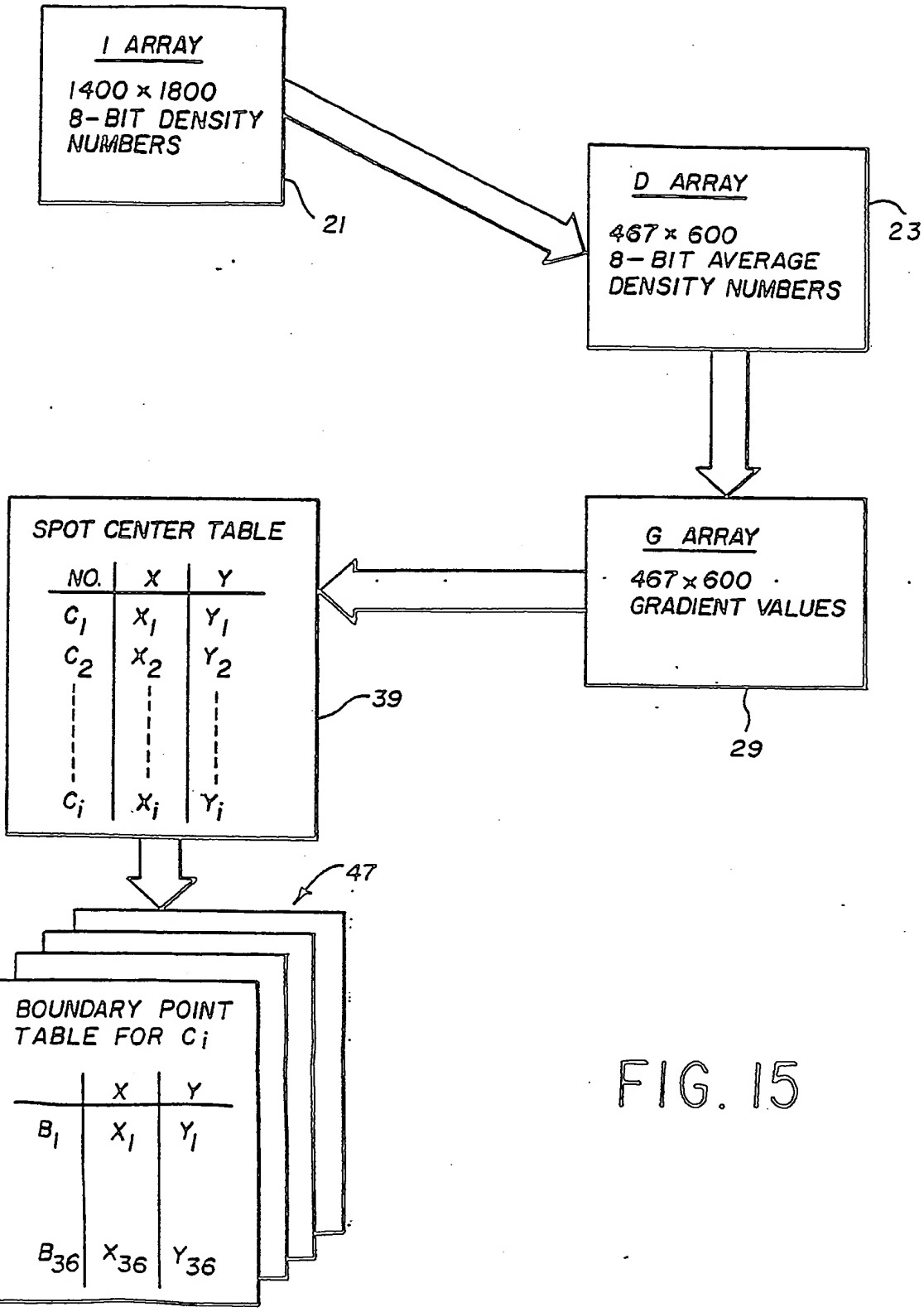
I ARRAY
1400 x 1800
8-BIT DENSITY
NUMBERS
21
D ARRAY
467 x 600
8-BIT AVERAGE
DENSITY NUMBERS
23
SPOT CENTER TABLE
NO. X Y
C1 X1 Y1
C2 X2 Y2
Ci Xi Yi
39
G ARRAY
467 x 600
GRADIENT VALUES
29
47
BOUNDARY POINT
TABLE FOR Ci
X Y
B1 X1 Y1
B36 X36 Y36

FIG. 15

INTERNATIONAL SEARCH REPORT

International Application No PCT/US86/00984

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) 3

According to International Patent Classification (IPC) or to both National Classification and IPC

INT. CL.4 G01N 27/26; C12Q 1/16; G01N 33/558
U.S. CL. 204/182.8; 435/35; 436/514

II. FIELDS SEARCHED

Minimum Documentation Searched 4

Classification System	Classification Symbols
U.S.	204/182.8;204/182.9;204/299R;435/35; See Attachment 436/15;436/149;436/164;436/514; See Attachment

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched 6

III. DOCUMENTS CONSIDERED TO BE RELEVANT 14

Category *	Citation of Document, 16 with indication, where appropriate, of the relevant passages 17	Relevant to Claim No. 18
Y	GARRISON, ET AL., J. Biol. Chem., (Baltimore, MD, USA), Vol. 257, No. 21, (1982), pages 13144-13149, see entire document.	1-6
Y	WALTON, ET AL., J. Biol. Chem., (Baltimore, MD, USA), Vol. 254, No. 16, (1979), pages 7951-7960, see pages 7951-7952 and 7958-7960.	2-6
A	GARRELS, J. I., J. Biol. Chem., (Baltimore, MD, USA), Vol. 254, No. 16, (1979), pages 7961-7977, see entire document.	1-6
A	BOSSINGER, ET AL., J. Biol Chem., (Baltimore, MD, USA), Vol. 254, No. 16, (1979), pages 7986-7998, see entire document.	1-6

* Special categories of cited documents: 15

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search 2	Date of Mailing of this International Search Report 2
05 August 1986	13 AUG 1986
International Searching Authority 1	**Signature of Authorized Officer 20**
ISA/US	Stephen C. Wieder

Form PCT/ISA/210 (second sheet) (October 1981)

PCT/US86/00984

ATTACHMENT

II. FIELDS SEARCHED

U.S. 435/68; 435/173; 435/808; 436/516; 436/805; 530/344; 935/85; 935/86; 935/87